HAMILTON, BROOK, SMITH & REYNOLDS, P.C.

A

09/23/99 jc520 U.S. PTO

UTILITY PATENT APPLICATION TRANSMITTAL

(Only for new nonprovisional applications under 37 C.F.R. 1.53(b))

Attorney Docket No.	WHI9721p3MC2
First Named Inventor or Application Identifier	Andreas Stahl
Express Mail Label No.	EL 192627573 US

09/23/99 jc675 U.S. PTO 09/405504

Title of Invention: FATTY ACID TRANSPORT PROTEINS

APPLICATION ELEMENTS

See MPEP chapter 600 concerning utility patent application contents.

ADDRESS TO: **Assistant Commissioner for Patents, Box Patent Application, Washington, D.C. 20231**

1. ☐ Fee Transmittal Form *(Submit an original, and a duplicate for fee processing)*
2. ☒ Specification [Total Pages 115] *(preferred arrangement set forth below)*
   - Descriptive title of the invention
   - Cross References to Related Applications
   - Statement Regarding Fed sponsored R & D
   - Reference to microfiche Appendix
   - Background of the Invention
   - Summary of the Invention
   - Brief Description of the Drawings
   - Detailed Description
   - Claim(s)
   - Abstract of the Disclosure
3. ☒ Drawing(s) (35 U.S.C. 113) [Total Sheets 122] ____Formal __X__Informal
4. ☐ Oath or Declaration/POA [Total Pages]
   a. ☐ Newly executed (original or copy)
   b. ☐ Copy from a prior application (37 C.F.R. 1.63(d)) *(for continuation/divisional with Box 17 completed)* **[NOTE Box 5 below]**
      i. ☐ DELETION OF INVENTOR(S) Signed statement attached deleting inventor(s) named in the prior application, see 37 C.F.R. 1.63(d)(2) and 1.33(b).
5. ☐ Incorporation By Reference *(useable if Box 4b is checked)* The entire disclosure of the prior application, from which a copy of the oath or declaration is supplied under Box 4b, is considered as being part of the disclosure of the accompanying application and is hereby incorporated by reference therein.
6. ☐ Microfiche Computer Program *(Appendix)*
7. ☐ Nucleotide and/or Amino Acid Sequence Submission (if applicable, all necessary)
   a. ☐ Computer Readable Copy
   b. ☐ Paper Copy (identical to computer copy)
   c. ☐ Statement verifying identity of above copies

ACCOMPANYING APPLICATION PARTS

8. ☐ Assignment Papers (cover sheet & documents)
9. ☐ 37 C.F.R. 3.73(b) Statement ☐ Power of Attorney *(when there is an assignee)*
10. ☐ English Translation Document *(if applicable)*
11. ☐ Information Disclosure Statement (IDS)/PTO-1449 ☐ Copies of IDS Citations
12. ☐ Preliminary Amendment
13. ☒ Return Receipt Postcard (MPEP 503) (2) *(Should be specifically itemized)*
14. ☐ Small Entity Statement(s) ☐ Statement filed in prior application, Status still proper and desired
15. ☐ Certified Copy of Priority Document(s) *(if foreign priority is claimed)*
16. ☐ Other:.....

17. **If a CONTINUING APPLICATION,** check appropriate box and supply the requisite information:
    ☐ Continuation ☐ Divisional ☒ Continuation-in-part (CIP) of prior application No.: 09 / 232,201
    Prior application information: Examiner: _____ Group Art Unit: 1636

18. CORRESPONDENCE ADDRESS

NAME	Doreen M. Hogle, Esq.				
	HAMILTON, BROOK, SMITH & REYNOLDS, P.C.				
ADDRESS	Two Militia Drive				
CITY	Lexington	STATE	MA	ZIP CODE	02421-4799
COUNTRY	USA	TELEPHONE	(781) 861-6240	FAX	(781) 861-9540

Signature	Carol A. Egner	Date	September 23, 1999
Submitted by Typed or Printed Name	Carol A. Egner	Reg. Number	38,866

Rev January 12, 1999

Date: 23 Sep 99 Express Mail Label No. EL 192627573 US

Inventors: Andreas Stahl, David J. Hirsch, Harvey F. Lodish, Ruth E. Gimeno and Louis A. Tartaglia

Attorney's Docket No.: WHI97-21p3MC2

FATTY ACID TRANSPORT PROTEINS

RELATED APPLICATIONS

This application is a continuation-in-part of U.S. Patent Application Number 09/232,201 filed January 14, 1999, which claims the benefit of U.S. Provisional Application No. 60/110,941 filed December 4, 1998; U.S. Provisional Application No. 60/093,491 filed July 20, 1998; and U.S. Provisional Application No. 60/071,374 filed January 15, 1998. The teachings of each of these referenced applications are incorporated herein by reference in their entirety.

GOVERNMENT SUPPORT

The invention was supported, in whole or in part, by a grant from the National Heart, Lung, and Blood Institute (HL41484), by National Institutes of Health Grant DK 47618 and National Institutes of Health Grant 5 T32 CA 09541. The United States Government has certain rights in the invention.

BACKGROUND OF THE INVENTION

Long chain fatty acids (LCFAs) are an important source of energy for most organisms. They also function as blood hormones, regulating key metabolic functions such as hepatic glucose production. Although LCFAs can diffuse through the hydrophobic core of the plasma membrane into cells, this nonspecific transport cannot account for the high affinity and specific transport of LCFAs exhibited by cells such as

cardiac muscle, hepatocytes, enterocytes, and adipocytes. The molecular mechanisms of LCFA transport remains largely unknown. Identifying these mechanisms can lead to pharmaceuticals that modulate fatty acid uptake by the intestine and by other organs, thereby alleviating certain medical conditions (e.g. obesity).

SUMMARY OF THE INVENTION

Described herein is a diverse family of fatty acid transport proteins (FATPs) which are evolutionarily conserved; these FATPs are plasma membrane proteins which mediate transport of LCFAs across the membranes and into cells. Members of the FATP family described herein are present in a wide variety of organisms, from mycobacteria to humans, and exhibit very different expression patterns in tissues among the organisms. FATP family members are expressed in prokaryotic and eukaryotic organisms and comprise characteristic amino acid domains or sequences which are highly conserved across family members. In addition, the function of the FATP gene family is conserved throughout evolution, as shown by the fact that the *Caenorhabditis (C). elegans* and mycobacterial FATPs described herein facilitate LCFA uptake when they are overexpressed in COS cells or *Escherichia (E.) coli,* respectively. FATPs are expressed in a wide variety of tissues, including all tissues which are important to fatty acid metabolism (uptake and processing).

In specific embodiments, FATPs of the present invention are from such diverse organisms as humans *(Homo (H.) sapiens)*, mice, (*Mus* (M.) *musculus*), *F. rubripes, C. elegans, Drosophila (D.) melanogaster, Saccharomyces (S.) cerevisiae, Aspergillus nidulans, Cochliobolu heterostrophus, Magnaporthe grisea* and *Mycobacterium (M.),* such as *M. tuberculosis.* As described herein, four novel mouse FATPs, referred to as mmFATP2, mmFATP3, mmFATP4 and mmFATP5, and six human FATPs, referred to as hsFATP1, hsFATP2, hsFATP3, hsFATP4, hsFATP5 and hsFATP6, have been identified. All four novel murine FATPs (mmFATP2-5) and a previously identified murine FATP (renamed herein FATP1) have orthologs in humans (hsFATP1-5); the

sixth human FATP (hsFATP6) does not as yet have a mouse ortholog. The expression patterns of these FATPs vary, as described in detail below.

The present invention relates to FATP family members from prokaryotes and eukaryotes, nucleic acids (DNA, RNA) encoding FATPs, and nucleic acids which are useful as probes or primers (e.g., for use in hybridization methods, amplification methods) for example, in methods of detecting FATP-encoding genes, producing FATPs, and purifying or isolating FATP-encoding DNA or RNA. Also the subject of this invention are antibodies (polyclonal or monoclonal) which bind an FATP or FATPs; methods of identifying additional FATP family members (for example, orthologs of those FATPs described herein by amino acid sequence) and variant alleles of known FATP genes; methods of identifying compounds which bind to an FATP, or modulate or alter (enhance or inhibit) FATP function; compounds which modulate or alter FATP function; methods of modulating or altering (enhancing or inhibiting) FATP function and, thus, LCFA uptake into tissues of a mammal (e.g. human) by administering a compound or molecule (a drug or agent) which increases or reduces FATP activity; and methods of targeting compounds to tissues by administering a complex of the compound to be targeted to tissues and a component which is bound by an FATP present on cells of the tissues to which the compound is to be targeted. For example, a complex of a drug to be delivered to the liver and a component which is bound by an FATP present on liver cells (e.g., FATP5) can be administered.

In one embodiment, the present invention relates to modulating or altering (enhancing or inhibiting/reducing) LCFA uptake in the small intestine and, thus, increasing or reducing the number of calories in the form of fats available to an individual. In another embodiment, the present invention relates to inhibiting or reducing LCFA uptake in the small intestine in order to reduce circulating fatty acid levels; that is, LCFA uptake in the small intestine is reduced and, therefore, circulating (blood) levels are not as high as they otherwise would be. FATP4 has been shown to be expressed in epithelial cells of the small intestine and particularly in the brush border layer of the small intestine. FATP2 has also been shown to be expressed at low levels

in epithelial cells of the small intestine, particularly in the duodenum. In contrast, FATP1, FATP3, FATP5 and FATP6 were not detected in any of the intestinal tissues. Thus, also described herein are FATPs which are present in the epithelial cell layer of the small intestine where they mediate LCFA uptake. These FATPs, particularly FATP4 and also FATP2, are targets for methods and drugs which block their function or activity and are useful in treating obesity, diabetes and heart disease. The ability of these FATPs to mediate fat uptake can be modulated or altered (enhanced or inhibited), thus modulating fat uptake in the small intestine. This can be done, for example, by administering to an individual, such as a human or other animal, a drug which blocks interaction of LCFAs with FATP4 and/or FATP2 in the small intestine, thus inhibiting LCFA passage into the cells of the small intestine. As a result, fat absorption is reduced and, although the individual has consumed a certain quantity of fat, the LCFAs are not absorbed to the same extent they would have been in the absence of the compound administered.

Thus, one embodiment of this invention is a method of reducing LCFA uptake (absorption) in the small intestine and, as a result, reducing caloric uptake in the form of fat. A further embodiment is a compound (drug) useful in inhibiting or reducing fat absorption in the small intestine. In another embodiment, the invention is a method of reducing circulating fatty acid levels by administering to an individual a compound which blocks interactions of LCFAs with FATP4 and/or FATP2 in the small intestine, thus inhibiting LCFA passage into cells of the small intestine. As a result, fatty acids pass into the circulatory system at a diminished level and/or rate, and circulating fatty acid levels are lower than they would be in the absence of the compound administered. This method is particularly useful for therapy in individuals who are at risk for or have hyperlipidemia. That is, it can be used to prevent the occurrence of elevated levels of lipids in the blood or to treat an individual in whom blood lipid levels are elevated. Also the subject of this invention is a method of identifying compounds which alter FATP function (and thus, in the case of FATP2 and/or FATP4, alter LCFA uptake in the small intestine).

In another embodiment, the present invention relates to a method of modulating or altering (enhancing or inhibiting) the function of FATP6, which is expressed at high levels in the heart. A method of inhibiting FATP6 function is useful, for example, in individuals with heart disease, such as ischemia, since reducing LCFA uptake into heart muscle in an individual who has ischemic heart disease, which may be manifested by, for example, angina or heart attack, can reduce symptoms or reduce the extent of damage caused by the ischemia. In this embodiment, a drug which inhibits FATP6 function is administered to an individual who has had or is having a heart attack, to reduce LCFA uptake by the individual's heart and, as a result, reduce the damage caused by ischemia. In a further embodiment, this invention is a method of targeting a compound, such as a therapeutic drug or an imaging reagent, to heart tissue by administering to an individual (e.g., a human) a complex of the compound and a component (e.g., a LCFA or LCFA-like compound) which is bound by an FATP (e.g., FATP6) present in cells of heart tissue.

In a further embodiment, LCFA uptake by the liver is modulated or altered (enhanced or reduced), in an individual. For example, a drug which inhibits the function of an FATP present in liver (e.g., FATP5) is administered to an individual who is diabetic, in order to reduce LCFA uptake by liver cells and, thus reduce insulin resistance.

The present invention, thus, provides methods which are useful to alter, particularly reduce, LCFA uptake in individuals and, as a result, to alter (particularly reduce), availability of the LCFAs for further metabolism. In a specific embodiment, the present invention provides methods useful to reduce LCFA uptake and, thus, fatty acid metabolism in individuals, with the result that caloric availability from fats is reduced, and circulating fatty acid levels are lower than they otherwise would be. These methods are useful, for example, as a means of weight control in individuals, (e.g., humans) and as a means of preventing elevated serum lipid levels or reducing serum lipid levels in humans. FATPs expressed in the small intestine, such as FATP4, are useful targets to be blocked in treating obesity (e.g., chronic obesity) or to be enhanced

in treating conditions in which enhanced LCFA uptake is desired (e.g., malabsorption syndrome or other wasting conditions).

The identification of this evolutionarily conserved fatty acid transporter family will allow a better understanding of the mechanisms whereby LCFAs traverse the lipid bilayer as well as yield insight into the control of energy homeostasis and its dysregulation in diseases such as diabetes and obesity.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows the amino acid sequence alignment of FATPs: mmFATP1 (SEQ ID NO:92), mmFATP2 (SEQ ID NO:93), mmFATP3 (SEQ ID NO:94), mmFATP4 (SEQ ID NO:95), mmFATP5 (SEQ ID NO:96), ceFATPa (SEQ ID NO:97), scFATP (SEQ ID NO:98) and mtFATP (SEQ ID NO:99). The underlining (amino acid residues 204-212 of mtFATP) indicates an AMP binding motif which is found in many classes of proteins; the underlining at amino acid residues 204-507 of the mtFATP sequence indicates the FATP 360 amino acid signature sequence.

Figures 2A-2D show results of LCFA uptake assays. Figures 2A-2D: COS cells were cotransfected using the DEAE-dextran method with the mammalian expression vectors pCDNA-CD2 either alone (control; Figure 2A) or in combination with one of the FATP-containing expression vectors (pCDNA-mmFATP1, Figure 2B; pCDNA-mmFATP2, Figure 2C; or pCMV-SPORT2-mmFATP5, Figure 2D) as described in Materials and Methods for Example 2. COS cells were gated on forward scatter (FSC) and side scatter (SS), and the results shown represent >10,000 cells. Cells exhibiting >300 CD2 fluorescence units (vertical line) representing 15% of all cells were deemed CD2 positive.

Figure 3 is a graph of fluorescence of cells expressing a FATP gene. As in Figures 2A-2D, COS cells were cotransfected with pCDNA-CD2 either alone (control) or in combination with one of the FATP-containing expression vectors (pCDNA-mmFATP1, pCDNA-mmFATP2, pCMV-SPORT2-mmFATP5, or pCDNA-ceFATPb). The mean BODIPY-FA fluorescence of the CD2-positive cells is plotted; results shown

represent the average of three experiments, each consisting of greater than 50,000 COS cells. Note that a logarithmic scale is used on the ordinate.

Figure 4 is a graph of the uptake of palmitate with time. The full-length coding region of mtFATP (squares) or a control protein (TFE3; circles) was subcloned into the inducible, prokaryotic expression vector pET (Novagen). Expression from the resulting plasmid was induced (solid symbols) in transformed *E. coli* cells with 1 mM isopropyl-β-D-thiogalactoside (IPTG) for 1 hour, or cells were left uninduced (open symbols). Data points were done in triplicate and counts were normalized to the number of bacteria as determined by OD_{600}.

Figure 5 is a phylogenetic tree produced by aligning complete and partial sequences for *FATP* genes from human, rat, mouse, puffer fish, *D. melanogaster*, *C. elegans*, *S. cerevisiae*, and *M. tuberculosis* using ClustalX and using these data to produce a phylogenetic tree using TreeViewPPC. The bar indicates the number of substitutions per residue, i.e., 0.1 corresponds to a distance of 10 substitutions per 100 residues.

Figure 6 shows a comparison of the FATP signature sequences of mmFATP1 (SEQ ID NO:1), mmFATP5, (SEQ ID NO:2), ceFATPa (SEQ ID NO:3), scFATP (SEQ ID NO:4) and mtFATP (SEQ ID NO:5).

Figure 7 shows the sequence identity among the FATP family members and VLACs, based on the 360 amino acid signature sequence of FATP from Figure 1.

Figures 8A and 8B are the mmFATP3 DNA sequence (SEQ ID NO:6).

Figure 9 is the mmFATP3 protein sequence (SEQ ID NO:7).

Figures 10A and 10B are the mmFATP4 DNA sequence (SEQ ID NO:8).

Figure 11 is the mmFATP4 protein sequence (SEQ ID NO:9).

Figures 12A and 12B are the mmFATP5 DNA sequence (SEQ ID NO:10).

Figure 13 is the mmFATP5 protein sequence (SEQ ID NO:11).

Figures 14A and 14B are the hsFATP2 DNA sequence (SEQ ID NO:12).

Figure 15 is the hsFATP2 protein sequence (SEQ ID NO:13).

Figures 16A and 16B are the hsFATP3 DNA sequence (SEQ ID NO:14).

Figure 17 is the hsFATP3 protein sequence (SEQ ID NO:15).

Figures 18A and 18B are the hsFATP4 DNA sequence (SEQ ID NO:16).

Figure 19 is the hsFATP4 protein sequence (SEQ ID NO:17).

Figures 20A and 20B are the hsFATP5 DNA sequence (SEQ ID NO:18).

Figure 21 is the hsFATP5 protein sequence (SEQ ID NO:19).

Figures 22A and 22B are the hsFATP6 DNA sequence (SEQ ID NO:20).

Figure 23 is the hsFATP6 protein sequence (SEQ ID NO:21).

Figures 24A and 24B are the mtFATP DNA sequence (SEQ ID NO:22).

Figure 25 is the mtFATP protein sequence (SEQ ID NO:23).

Figure 26 shows the DNA sequence (SEQ ID NO:24) and predicted amino acid sequence (SEQ ID NO:25) of human FATP1.

Figure 27 shows the DNA sequence (SEQ ID NO:26) and predicted amino acid sequence (SEQ ID NO:27) of human FATP4.

Figure 28A is a hydrophobicity plot for hsFATP1, showing that it has multiple membrane-spanning domains.

Figure 28B is the amino acid composition of hsFATP1.

Figure 28C is a hydrophilicity plot for hsFATP1, made using the Kyte-Doolittle method, averaging hydrophilicity values for 18 amino acid residues at a time.

Figure 29A is a hydrophobicity plot for hsFATP4, showing that it has multiple membrane-spanning domains.

Figure 29B is a listing of the amino acid composition of hsFATP4.

Figure 29C is a hydrophilicity plot for hsFATP4, made using the Kyte-Doolittle method, averaging hydrophilicity values for 18 amino acid residues at a time.

Figures 30A and 30B show a comparison of the nucleotide sequence of human FATP1 (SEQ ID NO:28) and the nucleotide sequence of mouse FATP1 (SEQ ID NO:29).

Figures 31A and 31B show a comparison of the nucleotide sequence of human FATP4 (SEQ ID NO:30) and the nucleotide sequence of mouse FATP4 (SEQ ID NO:31).

Figure 32 shows a comparison of the amino acid sequence of human FATP1 (SEQ ID NO:32) and the amino acid sequence of mouse FATP1 (SEQ ID NO:33). Shaded amino acid residues match the concensus sequence exactly

Figure 33 shows a comparison at the amino acid level of human FATP4 (SEQ ID NO:34) and mouse FATP4 (SEQ ID NO:35). Shaded amino acid residues match the concensus sequence exactly.

Figure 34 shows the nucleotide sequence (SEQ ID NO:36) and predicted amino acid sequence (SEQ ID NO:37) of hsFATP6.

Figure 35A is a hydrophobicity plot for hsFATP6, showing that it has multiple membrane-spanning domains.

Figure 35B is a listing of the amino acid composition of hsFATP6.

Figure 35C is a hydrophilicity plot for hsFATP6, made using the Kyte-Doolittle method, averaging hydrophilicity values for 18 amino acid residues at a time.

Figure 36 shows an alignment of the amino acid sequences of hsFATP1 (SEQ ID NO:38), hsFATP4 (SEQ ID NO:39) and hsFATP6 (SEQ ID NO:40). Shaded amino acid residues match the concensus sequence exactly.

Figure 37 shows results of assessment of fatty acid uptake by human FATP1 and human FATP4. The percent of CD2-positive cells exhibiting a BODIPY-fluorescence of more than 300 arbitrary units is plotted for the three different conditions tested.

Figure 38 is a graph showing uptake of tritiated oleate, with time, by 293 cells transfected with either (diamonds) a plasmid for expression of human FATP4 or (squares) a control plasmid.

Figure 39 is an illustration of the amino acid sequences of human FATP4 (SEQ ID NO:41) and mouse FATP4 (SEQ ID NO:42) compared to human FATP1 (SEQ ID NO:43). Shown by underlining are the FATP consensus sequence (236-556 of hsFATP1) and the AMP-binding motif (246-254 of hsFATP1). The human FATPs were cloned by screening libraries with sequences from ESTs (expressed sequence tags). Mouse FATP4 was cloned by PCR using degenerate primers.

Figure 40 is a graph showing the uptake, with time, of tritiated oleate by mouse enterocytes in the presence of no oligonucleotide (squares), sense oligonucleotide (circles) or antisense oligonucleotide (diamonds).

Figure 41 is a bar graph showing uptake of tritiated oleate, by mouse enterocytes in the presence of various concentrations of antisense (solid bars), mismatch (stippled bars) or sense (lined bars) oligonucleotides.

Figure 42 is a bar graph showing uptake of tritiated oleate and uptake of ^{35}S-labeled methionine by mouse enterocytes to which were added no oligonucleotide, the antisense oligonucleotide, or the mismatch oligonucleotide.

Figure 43A is the nucleotide sequence of the gene encoding mouse FATP4 (SEQ ID NO:44).

Figure 43B is the amino acid sequence of mouse FATP4 protein (SEQ ID NO:45).

Figures 44A, 44B, and 44C are the hsFATP1 DNA sequence (SEQ ID NO:46). Coding region: 175-2115 (1941 nt).

Figure 45 is the hsFATP1 protein sequence (SEQ ID NO:47).

Figures 46A and 46B are the hsFATP2 DNA sequence (SEQ ID NO:48). Coding region: 223-2085 (1863 nt).

Figure 47 is the hsFATP2 protein sequence (SEQ ID NO:49).

Figure 48 is the partial DNA sequence of hsFATP3 (SEQ ID NO:50). Coding region: 1-993.

Figure 49 is the partial protein sequence of hsFATP3 (SEQ ID NO:51).

Figures 50A, 50B, and 50C are the hsFATP4 DNA sequence (SEQ ID NO:52). Coding region: 208-2139 (1932 nt).

Figure 51 is the hsFATP4 protein sequence (SEQ ID NO:53).

Figure 52 is the hsFATP5 partial DNA sequence (SEQ ID NO:54). Coding region: 1-1062.

Figure 53 is the hsFATP5 partial protein sequence (SEQ ID NO:55).

Figures 54A, 54B, and 54C are the hsFATP6 DNA sequence (SEQ ID NO:56). Coding region: 643-2502 (1860 nt).

Figure 55 is the hsFATP6 protein sequence (SEQ ID NO:57).

Figures 56A, 56B, and 56C are the rnFATP1 DNA sequence (rn=*Rattus norvegicus*; (SEQ ID NO:58). Coding region: 75-2015 (1941 nt).

Figure 57 is the rnFATP1 protein sequence (SEQ ID NO:59).

Figure 58A, 58B, and 58C are the rnFATP2 DNA sequence (SEQ ID NO:60). Coding region: 795-2657 (1863 nt).

Figure 59 is the rnFATP2 protein sequence (SEQ ID NO:61).

Figure 60A and 60B are the rnFATP4 partial DNA sequence (SEQ ID NO:62). Coding region: 1-1218.

Figure 61 is the rnFATP4 partial DNA sequence (SEQ ID NO:63).

Figure 62A, 62B, and 62C are the mmFATP1 DNA sequence (SEQ ID NO:64). Coding region: 1-1944.

Figure 63 is the mmFATP1 protein sequence (SEQ ID NO:65).

Figures 64A and 64B are the mmFATP2 DNA sequence (SEQ ID NO:66). Coding region: 121-1992 (1872 nt).

Figure 65 is the mmFATP2 protein sequence (SEQ ID NO:67).

Figures 66A and 66B are the mmFATP3 partial DNA sequence (SEQ ID NO:68). Coding region: 1-1830.

Figure 67 is the mmFATP3 partial protein sequence (SEQ ID NO:69).

Figures 68A, 68B, and 68C are the mmFATP4 DNA sequence (SEQ ID NO:70). Coding region: 1-1932.

Figures 69 is the mmFATP4 protein sequence (SEQ ID NO:71).

Figures 70A and 70B are the mmFATP5 DNA sequence (SEQ ID NO:72). Coding region: 60-2129.

Figure 71 is the mmFATP5 protein sequence (SEQ ID NO:73).

Figures 72A and 72B are the dmFATP partial DNA sequence (dm=*Drosophila melanogaster*; SEQ ID NO:74). Coding region: 1-1773.

Figures 73 is the dmFATP partial protein sequence (SEQ ID NO:75).

Figure 74 is the drFATP partial DNA sequence (dr=*Danio rerio*, zebrafish; SEQ ID NO:76) Coding region: 1-173.

Figure 75 is the drFATP partial protein sequence (SEQ ID NO:77).

Figure 76A and 76B are the ceFATPa DNA sequence (SEQ ID NO:78). Coding region: 1-1953.

Figure 77 is the ceFATPa protein sequence (SEQ ID NO:79).

Figures 78A and 78B are the ceFATPb DNA sequence (SEQ ID NO:80). Coding region: 1-1968.

Figure 79 is the ceFATPb protein sequence (SEQ ID NO:81).

Figures 80A and 80B are the chFATP DNA sequence (SEQ ID NO:82; ch=*Cochliobolu heterostrophus*). Coding region: 1-1932.

Figure 81 is the chFATP protein sequence (SEQ ID NO:83).

Figure 82 is the anFATP partial protein sequence (*an=Aspergillus nidulans;* SEQ ID NO:84). Coding region: 1-597.

Figure 83 is the anFATP partial protein sequence (SEQ ID NO:85).

Figure 84 is the mgFATP partial DNA sequence (mg= *Magnaporthe grisea*, rice blast; SEQ ID NO:86). Coding region: 1-522.

Figure 85 is the mgFATP partial protein sequence (SEQ ID NO:87).

Figures 86A and 86B are the scFATP DNA sequence (SEQ ID NO:88). Coding region: 1-1872.

Figure 87 is the scFATP protein sequence (SEQ ID NO:89).

Figures 88A and 88B are the mtFATP DNA sequence (SEQ ID NO:90).

Figure 89 is the mtFATP protein sequence (SEQ ID NO:91). Coding region: 1-1794.

Figure 90 is a concensus sequence of the FATP signature sequence (SEQ ID NO:100), based on 23 independent sequences aligned in ClustalX. The height of the bar at each amino acid residue position indicates the degree of conservation at that position. Gaps have been inserted to maintain the strength of the alignment.

Figure 91 is a hydrophilicity plot for hsFATP2, made using the Kyte-Doolittle method, averaging hydrophilicity values for 18 amino acid residues at a time.

Figure 92 is a hydrophilicity plot for the hsFATP3 partial protein, made using the Kyte-Doolittle method, averaging hydrophilicity values for 18 amino acid residues at a time.

Figure 93 is a hydrophilicity plot for the hsFATP5 partial protein, made using the Kyte-Doolittle method, averaging hydrophilicity values for 18 amino acid residues at a time.

Figures 94A and 94B are a representation of the DNA sequence (SEQ ID NO:101) of the hsFATP3 gene, and the amino acid sequence (SEQ ID NO:102) of the hsFATP3 protein.

Figure 95. Mammalian expression constructs containing either hsFATP4 (squares and triangles) or empty control vector (circles) were stably transfected into 293 cells. Short-term uptake of Bodipy-FA in the presence of BSA was determined by FACS. The mean fluorescence of the viable cell population is expressed in arbitrary fluorescence units. FATP4 protein expression was determined by densitometry of anti-FATP4 Western blots, and is expressed in arbitrary units.

Figure 96. Short-term uptake of Bodipy-palmitate (1 μM), either by control cells (black bars) or FATP4-expressing cells (hatched bars), was measured in the presence of 0, 10, 100 μM unlabeled palmitate. FA uptake was quantified by FACS and expressed in arbitrary fluorescence units.

Figure 97. The rate of [^{2}H]palmitate uptake by 293 cells, which were stably transfected with a construct for either human FATP4 (diamonds) or an empty vector (circles), was compared to that of isolated enterocytes (squares).

Figure 98. Isolated enterocytes were incubated for 48h with increasing concentrations of the FATP4 antisense oligonucleotide or with 100 μM of a randomized control oligonucleotide with identical nucleotide composition to the FATP4 antisense oligonucleotide. The uptake of oleate by the enterocytes was then measured over a 5 min time interval (solid bars). In parallel, the levels of FATP4 protein and, as a loading

control, β-catenin, were determined by Western blotting and quantitated using densitometry (hatched bars). FA uptake and FATP4 protein levels were normalized to that of untreated cells. The averages and standard deviations of 4 independent experiments are shown.

Figure 99. Uptake rates of [^{3}H]oleate, [^{3}H]palmitate and [^{35}S]methionine by primary enterocytes were measured after 48h incubation with either 100 μM FATP4 antisense (solid bars) or 100 μM randomized control oligonucleotide (hatched bars) and expressed as % of untreated cells.

DETAILED DESCRIPTION OF THE INVENTION

As described herein, FATPs are a large evolutionarily conserved family of proteins that mediate the transport of LCFAs into cells. The family includes proteins which are conserved from mycobacteria to humans and exhibit very different expression patterns in tissues. Specific embodiments described include FATPs from mice, humans, nematodes, fungi and mycobacteria which have been shown to be functional LCFA transporters. The term "fatty acid transport proteins" ("FATPs") as used herein, refers to the proteins described herein as FATP1, FATP2, FATP3, FATP4, FATP5 and FATP6, which have been described in one or more species of mammals, as well as mtFATP, ceFATP, scFATP, anFATP, mgFATP, and chFATP, and other proteins sharing at least about 50% amino acid sequence similarity, preferably at least about 60% sequence similarity, more preferably at least about 70% sequence similarity, and still more preferably, at least about 80% sequence similarity, and most preferably, at least about 90% sequence similarity in the approximately 360 amino acid signature sequence. The approximaely 360 amino acid FATP signature sequence is shown in Figure 1. The concensus sequence of the signature sequence is shown in Figure 90. The nomenclature used herein to refer to FATPs includes a species-specific prefix (e.g., mm, *Mus musculus*; hs or h, *Homo sapiens* or human; mt *M. tuberculosis*; dm. *D. melanogaster*; ce, *C. elegans*; sc, *Saccharomyces cerevisiae*) and a number such that mammalian homologues in different species share the same number. For example, six human and

five mouse *FATP* genes which are expressed in a variety of tissues are described herein and are referred to, respectively, as hsFATP1-hsFATP6 and mmFATP1-mmFATP5; for example, hsFATP4 and mmFATP4 are the human and mouse orthologs.

Expression patterns of human and mouse FATPs have been assessed and are described below. Briefly, results of these assessments show that FATP5 is a liver-specific gene. FATP2 is highly expressed in liver and kidney. Both of these proteins, as well as FATP4 and FATPs from nematodes and mycobacteria, have been shown to be functional LCFA transporters. Results have also shown that FATP4 mRNA is present at high levels in epithelial cells of two regions of the small intestine (the jejunum and ileum) and at lower, but significant, levels in a third region (the duodenum). They further showed that FATP2 mRNA is present in epithelial cells of the duodenum at a level similar to that of FATP4 mRNA levels, but is present at lower levels in the jejunum and ileum. FATP4 mRNA was absent from other cell types of the small intestine and no FATP4 mRNA could be detected in any cells of the colon. No signals above background could be detected for FATP1, FATP3 and FATP5 in any of the intestinal tissues. Thus, FATP4 is the major FATP in the mouse small intestine, which supports a major role for FATP4 (along with FATP2 to a lesser extent) in absorption of free fatty acids. hsFATP4 was clearly expressed in the jejunum and ileum; expression was absent in the stomach. This, too, is consistent with a major role for FATP4 in absorption of fatty acids in the human gut. Analysis of FATP expression in human tissues, also described in detail below, showed that hsFATP6, which has no mouse ortholog as yet, is expressed at high levels in the heart and at low levels in the placenta, but is undetectable in the other tissues assessed (Example 9). This is consistent with a major role for FATP6 in absorption of fatty acids in the heart.

Long chain fatty acids (LCFAs) are an important energy source for pro- and eukaryotes and are involved in diverse cellular processes, such as membrane synthesis, intracellular signaling, protein modification, and transcriptional regulation. In developed Western countries, human dietary lipids are mainly di- and triglycerides and account for approximately 40% of caloric intake (Weisburger, J. H. (1997) *J. Am. Diet.*

Assoc. 97:S16-S23). These lipids are broken down into fatty acids and glycerol by pancreatic lipases in the small intestine (Chapus, C., Rovery, M., Sarda, L & Verger, R. (1988) *Biochimie 70*:1223-34); LCFAs are then transported into brush border cells, where the majority is re-esterified and secreted into the lymphatic system as chylomicrons (Green, P.H. & Riley, J.W. (1981) *Aust. N.Z.J. Med. 11*:84-90). Fatty acids are liberated from lipoproteins by the enzyme lipoprotein lipase, which is bound to the luminal side of endothelial cells (Scow, R.O. & Blachette-Mackie, E.J. (1992) *Mol. Cell. Biochem 116*:181-191). "Free" fatty acids in the circulation are bound to serum albumin (Spector, A.A. (1984) *Clin. Physiol. Biochem 2*:123-134) and are rapidly incorporated by adipocytes, hepatocytes, and cardiac muscle cells. The latter derive 60-90% of their energy through the oxidation of LCFAs (Neely, J.F. Rovetto, M.J. & Oram, J.F. (1972) *Prog. Cardiovasc. Dis: 15*:289-329). Although saturable and specific uptake of LCFAs has been demonstrated for intestinal cells, hepatocytes, cardiac myocytes, and adipocytes, the molecular mechanisms of LCFA transport across the plasma membrane have remained controversial (Hui, T.Y. & Bernlohr, D.A. (1997) *Front. Biosci. 15*:d222-31-d231; Schaffer, J.E. & Lodish, H.F, (1995) *Trends Cardiovasc. Med. 5*:218-224). Described herein is a large family of highly homologous mammalian LCFA transporters which show wide expression, including in all tissues relevant to fatty acid metabolism. Further described are novel members of this family in other species, including mycobacterial and nematode FATPs which, like their mammalian counterparts, are functional fatty acid transporters.

The discovery of a diverse but highly homologous family of FATPs is reminiscent of the glucose transporter family. In a manner similar to the FATPs, the glucose transporters have very divergent patterns of tissue expression (McGowan, K.M., Long, S.D. & Pekala, P.H. (1995) *Pharmacol. Ther. 66*:465-505). The FATPs, like glucose transporters, may also differ in their substrate specificities, uptake kinetics, and hormonal regulation (Thorens, B. (1996) *Am. J. Physiol. 270*:G541-G553). Indeed, the levels of fatty acids in the blood, like those of glucose, can be regulated by insulin and are dysregulated in diseases such as noninsulin-dependent diabetes and obesity (Boden,

G. (1997) *Diabetes 46*:3-10). The underlying mechanisms for the regulation of free fatty acid concentrations in the blood are not understood, but could be explained by hormonal modulation of FATPs.

Insulin-resistance is thought to be the major defect in non insulin-dependent diabetes mellitus (NIDDM) and is one of the earliest manifestations of NIDDM (McGarry (1992) *Science* 258:766-770). Free fatty acids (FFAs) may provide an explanation for why obesity is a risk factor for NIDDM. Plasma levels of FFAs are elevated in diabetic patients (Reaven *et al.* (1988) *Diabetes* 37:1020). Elevated plasma free fatty acids (FFAs) have been demonstrated to induce insulin-resistance in whole animals and humans (Boden (1998) *Front. Biosci.* 3:D169-D175). This insulin-resistance is likely mediated by effects of FFAs on a variety of issues. FFAs added to adipocytes *in vitro* induce insulin resistance in this cell type as evidenced by inhibition of insulin-induced glucose transport (Van Epps-Fung *et al.* (1997) *Endocrinology* 138:4338-4345). Rats fed a high fat diet developed skeletal muscle insulin resistance as evidenced by a decrease in insulin-induced glucose uptake by skeletal muscle (Han *et al.*, (1997) *Diabetes* 46:1761-1767). In addition, elevated plasma FFAs increase insulin-suppressed endogenous glucose production in the liver (Boden (1998) *Front. Biosci.* 3:D169-D175), thus increasing hepatic glucose output. It has been postulated that the adverse effects of plasma free fatty acids are due to the FFAs being taken up into the cell, leading to an increase in intracellular long chain fatty acyl CoA; intracellular long chain acyl CoAs are thought to mediate the effects of FFAs inside the cell. Thus, fatty acid induced insulin-resistance may be prevented by blocking uptake of FFAs into select tissues, in particular liver (by blocking FATP2 and/or FATP5), adipocyte (by blocking FATP1), and skeletal muscle (by blocking FATP1). Blocking intestinal fat absorption (by blocking FATP4) is also expected to reduce plasma FFA levels and thus improve insulin resistance.

During the pathogenesis of NIDDM insulin-resistance can initially be counteracted by increasing insulin output by the pancreatic beta cell. Ultimately, this compensation fails, beta cell function decreases and overt diabetes results (McGarry

(1992) *Science* 258: 766-770). Manipulating beta cell function is a second point where fatty acid transporter blockers may be beneficial for diabetes. While no FATP homolog has been identified so far that is expressed in the beta cell of the pancreas, the data described below suggest the existence of such a transporter and the sequence information included herein provides the means to identify such a transporter by degenerate PCR, using primers to regions conserved in all FATP family members or by low stringency hybridization. It has been demonstrated that exposure of pancreatic beta-cells to FFAs increases the basal rate of insulin secretion; this in turn leads to a decrease in the intracellular stores of insulin, resulting in decreased capacity for insulin secretion after chronic exposure (Bollheimer *et al.,* (1998) *J. Clin. Invest.* 101:1094-1101). The effects of FFAs are again likely to be mediated by intracellular long chain fatty acyl CoA molecules (Liu *et al.,* (1998) *J. Clin. Invest.* 101:1870-1875). FFAs have also been demonstrated to increase beta cell apoptosis (Shimabukuro *et al.,* (1998) *Proc. Nat. Acad. Sci. USA* 95:2498-2502), possibly contributing to the decrease in beta cell numbers in late stage NIDDM.

Another finding with potentially broad implications is the identification of a FATP homologue in *M. tuberculosis*. Tuberculosis causes more deaths worldwide than any other infectious agent and drug-resistant tuberculosis is re-emerging as a problem in industrialized nations (Bloom, B.R. & Small, P.M. (1998) *N. Engl. J. Med. 338*:677-678). *Mycobacterium tuberculosis* has about 250 enzymes involved in fatty acid metabolism, compared with only about 50 in *E. coli*. It has been suggested that, living as a pathogen, the mycobacteria are largely lipolytic, rather than lipogenic, relying on the lipds within mammalian cells and the tubercle (Cole, S.T. *et al., Nature 393*:537-544 (1998)). The *de novo* synthesis of fatty acids in *Mycobacterium leprae* is insufficient to maintain growth (Wheeler, P.R., Bulmer, K & Ratledge, C. (1990) *J. Gene. Microbiol. 136*:211-217). Thus, it is reasonable to expect that inhibitors of mtFATP will serve as therapeutics for tuberculosis. FATPs expressed in mycobacteria can be targeted to reduce or prevent replication of mycobacteria (e.g., to reduce or prevent replication of *M. tuberculosis*) and, thus, reduce or prevent their adverse effects.

For example, a FATP or FATPs expressed by *M. tuberculosis* can be targeted and inhibited, thus reducing or preventing growth of this pathogen (and tuberculosis in humans and other mammals). An inhibitor of an *M. tuberculosis* FATP can be identified, using methods described herein (e.g., expressing the FATP in an appropriate host cell, such as *E. coli* or COS cells; contacting the cells with an agent or drug to be assessed for its ability to inhibit the FATP and, as a result, mycobacterial growth, and assessing its effects on growth). A drug or agent identified in this manner can be further tested for its ability to inhibit a *M. tuberculosis* FATP and *M. tuberculosis* infection in an appropriate animal model or in humans. A method of inhibiting mycobacterial growth, particularly growth of *M. tuberculosis*, and compounds useful as drugs for doing so are also the subject of this invention.

An isolated polynucleotide encoding mtFATP, like other polynucleotides encoding FATPs of the FATP family, can be incorporated into vectors, nucleic acids of viruses, and other nucleic acid constructs that can be used in various types of host cells to produce mtFATP. This mtFATP can be used, as it appears on the surface of cells, or in various artificial membrane systems, to assess fatty acid transport function, to identify ligands and molecules that are modulators of fatty acid transport activity. Molecules found to be inhibitors of mtFATP function can be incorporated into pharmaceutical compositions to administer to a human for the treatment of tuberculosis.

Particular embodiments of the invention are polynucleotides encoding a FATP of *Cochliobolus (Helminthosporium) heterostrophus* or portions or variants thereof, the isolated or recombinantly produced FATP, methods for assessing whether an agent binds to the chFATP, and further methods for assessing the effect of an agent being tested for its ability to modulate fatty acid transport activity. *Cochliobolus heterostrophus* is an ascomycete that is the cause of southern corn leaf blight, an economically important threat to the corn crop in the United States. The related species *C. sativus* causes crown rot and common root rot in wheat and barley. One or more FATPs of *C. heterostrophus* can be targeted for the identification of an inhibitor of chFATP function, which can be then be used as an agent effective against infection of

plants by *C. heterostrophus* and related organisms. Methods described herein that were applied in studying the expression of a FATP gene and the function of the FATP in its natural site of expression or in a host cell, can be used in the study of the chFATP gene and protein.

Magnaporthe grisea (rice blast) is an economically important fungal pathogen of rice. Further embodiments of the invention are nucleic acid molecules encoding a FATP of *Magnaporthe grisea*, portions thereof, or variants thereof, isolated mgFATP, nucleic acid constructs, and engineered cells expressing mgFATP. Other aspects of the invention are assays to identify an agent which binds to mgFATP and assays to identify an agent which modulates the function of mgFATP in cells in which mgFATP is expressed or in artificial membrane systems. Agents identified as inhibiting mgFATP activity can be developed into anti-fungal agents to be used to treat rice infected with rice blast.

Caenorhabditis elegans is a nematode related to plant pathogens and human parasites. An isolated polynucleotide which encodes ceFATP, like other polynucleotides encoding FATPs of the FATP family described herein, can be incorporated into nucleic acid vectors and other constructs that can be used in various types of cells to produce ceFATP. ceFATP as it occurs in cells or as it can be isolated or incorporated into various artificial or reconstructed membrane systems, can be used to assess fatty acid transport, and to identify ligands and agents that modulate fatty acid transport activity. Agents found by such assays to be inhibitors of ceFATP activity can be incorporated into compositions for the treatment of diseases caused by genetically related organisms with a FATP of similar sensitivity to the agents.

Aspergillus nidulans is one of a family of fungal species that can infect humans. Further embodiments of the invention of the family of polynucleotides encoding FATPs are polynucleotides encoding a FATP of *Aspergillus nidulans*, and vectors and host cells that can be constructed to comprise such polynucleotides. Further embodiments are a polypeptide encoded by such polynucleotides, portions thereof having one or more functions characteristic of a FATP, and various methods. The methods include those

for identifying agents that bind to anFATP and those for assessing the effect of an agent being tested for its ability to modulate fatty acid transport activity. Those agents found to inhibit fatty acid transport function can be used in compositions as anti-fungal pharmaceuticals, or can be modified for greater effectiveness as a pharmaceutical.

One aspect of the invention relates to isolated nucleic acids that encode a FATP as described herein, such as those FATPs having an amino acid sequence in Figure 45 (SEQ ID NO:47), Figure 47 (SEQ ID NO:49), Figures 94A and 94B (SEQ ID NO:102), Figure 51 (SEQ ID NO:53), Figure 53 (SEQ ID NO:55), and Figure 55 (SEQ ID NO:57) and nucleic acids closely related thereto as described herein.

Using the information provided herein, such as a nucleic acid sequence set forth in Figures 44A-44C (SEQ ID NO:46), Figures 46A and 46B (SEQ ID NO:48), Figures 94A and 94B (SEQ ID NO:101), Figures 50A-50C (SEQ ID NO:52), Figure 52 (SEQ ID NO:54), and Figures 54A-54C (SEQ ID NO:56), a nucleic acid of the invention encoding a FATP polypeptide may be obtained using standard cloning and screening methods, such as those for cloning and sequencing cDNA library fragments, followed by obtaining a full length clone. For example, to obtain a nucleic acid of the invention, a library of clones of cDNA of human or other mammalian DNA can be probed with a labeled oligonucleotide, such as a radiolabeled oligonucleotide, preferably about 17 nucleotides or longer, derived from a partial sequence. Clones carrying DNA identical to that of the probe can then be distinguished using stringent (also, "high stringency") hybridization conditions. By sequencing the individual clones thus identified with sequencing primers designed from the original sequence it is then possible to extend the sequence in both directions to determine the full length sequence. Suitable techniques are described, for example, in *Current Protocols in Molecular Biology* (F.M. Ausubel et al, eds), containing supplements through Supplement 42, 1998, John Wiley and Sons, Inc., especially chapters 5, 6 and 7.

Embodiments of the invention include isolated nucleic acid molecules comprising any of the following nucleotide sequences: *1.*) a nucleotide sequence which encodes a protein comprising the amino acid sequence of hsFATP1 (SEQ ID

NO:47), the amino acid sequence of hsFATP2 (SEQ ID NO:49), the amino acid sequence of hsFATP3 (SEQ ID NO:51), the amino acid sequence of hsFATP4 (SEQ ID NO: 53), the amino acid sequence of hsFATP5 (SEQ ID NO:55) or the amino acid sequence of hsFATP6 (SEQ ID NO:57); *2.*) nucleotide sequences of hsFATP1, hsFATP2, hsFATP3, hsFATP4, hsFATP5, or hsFATP6 (SEQ ID NO:46, 48, 101, 52, 54, or 56, respectively); *3.*) a nucleotide sequence which is complementary to the nucleotide sequence of hsFATP1 (SEQ ID NO:46), hsFATP2 (SEQ ID NO:48), hsFATP3 (SEQ ID NO:101), hsFATP4 (SEQ ID NO:52), hsFATP5 (SEQ ID NO:54) or hsFATP6 (SEQ ID NO:56); *4.*) a nucleotide sequence which consists of the coding region of hsFATP1 (SEQ ID NO:46), the coding region of hsFATP2 (SEQ ID NO:48), the coding region of hsFATP3 (SEQ ID NO:101), the coding region of hsFATP4 (SEQ ID NO:52), the coding region of hsFATP5 (SEQ ID NO:54), or the coding region of hsFATP6 (SEQ ID NO:56).

The invention further relates to nucleic acids (nucleic acid molecules or polynucleotides) having nucleotide sequences identical over their entire length to those shown in the figures, for instance Figures 44A-44C (SEQ ID NO:46), Figures 46A and 46B (SEQ ID NO:48), Figures 94A and 94B (SEQ ID NO:101), Figures 50A-50C (SEQ ID NO:52), Figure 52 (SEQ ID NO:54), and Figures 54A-54C (SEQ ID NO:56). It further relates to DNA, which due to the degeneracy of the genetic code, encodes a FATP encoded by one of the FATP-encoding DNAs, whose amino acid sequence is provided herein. Also provided by the invention are nucleic acids having the coding sequences for the mature polypeptides or fragments in reading frame with other coding sequences, such as those encoding a leader or secretory sequence, a pre-, or pro- or prepro- protein sequence. The nucleic acids of the invention encompass nucleic acids that include a single continuous region or discontinuous regions encoding the polypeptide, together with additional regions, that may also contain coding or non-coding sequences. The nucleic acids may also contain non-coding sequences, including, for example, but not limited to, non-coding 5' and 3' sequences, such as the transcribed, non-translated sequences, termination signals, ribosome binding sites, sequences that

stabilize mRNA, introns, polyadenylation signals, and additional coding sequences which encode additional amino acids. For example, a marker sequence that facilitates purification of the fused polypeptide can be encoded. In certain embodiments of the invention, the marker sequence can be a hexa-histidine peptide, as provided in the pQE vector (Qiagen, Inc.) and described in Gentz *et al., Proc. Natl. Acad. Sci. USA 86*: 821-824 (1989), or an HA tag (Wilson *et al., Cell 37*: 767 (1984)), or a sequence encoding glutathione S-transferase of *Schistosoma japonicum* (vectors available from Pharmacia; see Smith, D.B. and Johnson K.S., *Gene 67*:31 (1988) and Kaelin, W.G. *et al., Cell 70*:351 (1992)). Nucleic acids of the invention also include, but are not limited to, nucleic acids comprising a structural gene and its naturally associated sequences that control gene expression.

The invention further relates to variants, including naturally-occurring allelic variants, of those nucleic acids described specifically herein by DNA sequence, that encode variants of such polypeptides as those having the amino acid sequences shown in Figure 45 (SEQ ID NO:47), Figure 47 (SEQ ID NO:49), Figures 94A and 94B (SEQ ID NO:102), Figure 51 (SEQ ID NO:53) Figure 53 (SEQ ID NO:55), or Figure 55 (SEQ ID NO:57). Such variants include nucleic acids encoding variants of the above-listed amino acid sequences, wherein those variants have several, such as 5 to 10, 1 to 5, or 3, 2 or 1 amino acids substituted, deleted, or added, in any combination. Variants include polynucleotides encoding polypeptides with at least 95% but less than 100% amino acid sequence identity to the polypeptides described herein by amino acid sequence. Variant polynucleotides hybridize, under low to high stringency conditions, to the alleles described herein by DNA sequence. In one embodiment, variants have silent substitutions, additions and deletions that do not alter the properties and activities of the FATP. Allelic variants of the polynucleotides encoding hsFATP1 (Figure 45; SEQ ID NO:47), hsFATP2 (Figure 47; SEQ ID NO:49), hsFATP3 (Figures 94A and 94B; SEQ ID NO:102), hsFATP4 (Figure 51; SEQ ID NO:53), Figure 53 (SEQ ID NO:55) and hsFATP6 (Figure 55; SEQ ID NO:57) will be identified as mapping to chromosomal locations listed for the corresponding wild type genes in Table 2 in Example 1.

Orthologous genes are gene loci in different species that are sufficiently similar to each other in their nucleotide sequences to suggest that they originated from a common ancestral gene. Orthologous genes arise when a lineage splits into two species, rather than when a gene is duplicated within a genome. Proteins that are orthologs are encoded by genes of two different species, wherein the genes are said to be orthologous.

The invention further relates to polynucleotides encoding polypeptides which are orthologous to those polypeptides having a specific amino acid sequence described herein, such as the amino acid sequences shown in Figure 45 (SEQ ID NO:47), Figure 47 (SEQ ID NO:49), Figures 94A and 94B (SEQ ID NO:102), Figure 51 (SEQ ID NO:53), Figure 53 (SEQ ID NO:55), or Figure 55 (SEQ ID NO:57). These polynucleotides, which can be called ortholog polynucleotides, encode orthologous polypeptides that can range in amino acid sequence identity to a reference amino acid sequence described herein, from about 65% to less than 100%, but preferably 70% to 80%, more preferably 80% to 90%, and still more preferably 90% to less than 100%. Orthologous polypeptides can also be those polypeptides that range in amino acid sequence similarity to a reference amino acid sequence described herein from about 75% to 100%, within the signature sequence. The amino acid sequence similarity between the signature sequences of orthologous polypeptides is preferably 80%, more preferably 90%, and still more preferably, 95%. The ortholog polynucleotides encode polypeptides that have similar functional characteristics (e.g., fatty acid transport activity) and similar tissue distribution, as appropriate to the organism from which the ortholog polynucleotides can be isolated.

Ortholog polynucleotides can be isolated from (e.g., by cloning or nucleic acid amplification methods) a great number of species, as shown by the sample of FATPs from evolutionarily divergent species described herein (see, e.g., Figures 44A-C through Figure 89). Ortholog polynucleotides corresponding to those in Figure 45 (SEQ ID NO:47), Figure 47 (SEQ ID NO:49), Figures 94A and 94B (SEQ ID NO:101), Figure 51 (SEQ ID NO:53), Figure 52 (SEQ ID NO:54) and Figure 55 (SEQ ID NO:57) are

those which can be isolated from mammals such as rat, dog, chimpanzee, monkey, baboon, pig, rabbit and guinea pig, for example.

Further variants that are fragments of the nucleic acids of the invention may be used to synthesize full-length nucleic acids of the invention, such as by use as primers in a polymerase chain reaction. As used herein, the term primer refers to a single-stranded oligonucleotide which acts as a point of initiation of template-directed DNA synthesis under appropriate conditions (e.g., in the presence of four different nucleoside triphosphates and an agent for polymerization, such as DNA or RNA polymerase or reverse transcriptase) in an appropriate buffer and at a suitable temperature. The appropriate length of a primer depends on the intended use of the primer, but typically ranges from 15 to 30 nucleotides. Short primer molecules generally require cooler temperatures to form sufficiently stable hybrid complexes with the template. A primer need not reflect the exact sequence of the template, but must be sufficiently complementary to hybridize with a template. The term primer site refers to the area of the target DNA to which a primer hybridizes. The term primer pair refers to a set of primers including a 5' (upstream) primer that hybridizes with the 5' end of the DNA sequence to be amplified and a 3' (downstream) primer that hybridizes with the complement of the 3' end of the sequence to be amplified.

Further embodiments of the invention are nucleic acids that are at least 80% identical over their entire length to a nucleic acid described herein, for example a nucleic acid having the nucleotide sequence in Figures 44A-44C (SEQ ID NO:46), Figures 46A-46B (SEQ ID NO:48), Figures 94A and 94B (SEQ ID NO:101), Figures 50A-50C (SEQ ID NO:52), Figure 52 (SEQ ID NO:54), and Figures 54A-54C (SEQ ID NO:56). Additional embodiments are nucleic acids, and the complements of such nucleic acids, having at least 90% nucleotide sequence identity to the above-described sequences, and nucleic acids having at least 95% nucleotide sequence identity. In preferred embodiments, DNA of the present invention has 97% nucleotide sequence identity, 98% nucleotide sequence identity, or at least 99% nucleotide sequence identity with the DNA whose sequences are presented herein.

Other embodiments of the invention are nucleic acids that are at least 80% identical in nucleotide sequence to a nucleic acid encoding a polypeptide having an amino acid sequence as set forth in Figure 45 (SEQ ID NO:47), Figure 47 (SEQ ID NO:49), Figures 94A and 94B (SEQ ID NO:102), Figure 51 (SEQ ID NO:53), Figure 53 (SEQ ID NO:55) or Figure 55 (SEQ ID NO:57), or as such amino acid sequences are set forth elsewhere herein, and nucleic acids that are complementary to such nucleic acids. Specific embodiments are nucleic acids having at least 90% nucleotide sequence identity to a nucleic acid encoding a polypeptide having an amino acid sequence as described in the list above, nucleic acids having at least 95% sequence identity, and nucleic acids having at least 97% sequence identity.

The terms "complementary" or "complementarity" as used herein, refer to the natural binding of polynucleotides under permissive salt and temperature conditions by base-pairing. Complementarity between two single-stranded molecules may be "partial" in which only some of the nucleic acids bind, or it may be complete when total complementarity exists between the single-stranded molecules (that is, when A-T and G-C base pairing is 100% complete). The degree of complementarity between nucleic acid strands has significant effects on the efficiency and strength of hybridization between nucleic acid strands. This is of particular importance in amplification reactions, which depend on binding between nucleic acid strands.

The invention further includes nucleic acids that hybridize to the above-described nucleic acids, especially those nucleic acids that hybridize under stringent hybridization conditions. "Stringent hybridization conditions" or "high stringency conditions" generally occur within a range from about T_m minus 5°C (5° C below the strand dissociation temperature or melting temperature (T_m) of the probe nucleic acid molecule) to about 20° C to 25° C below T_m. As will be understood by those of skill in the art, the stringency of hybridization may be altered in order to identify or detect molecules having identical or related polynucleotide sequences. An example of high stringency hybridization follows. Hybridization solution is (6x SSC/10 mM EDTA/0.5% SDS/5x Denhardt's solution/100 μg/ml sheared and denatured salmon

sperm DNA). Hybridization is at 64-65°C for 16 hours. The hybridized blot is washed two times with 2x SSC/0.5% SDS solution at room temperature for 15 minutes each, and two times with 0.2x SSC/0.5% SDS at 65°C, for one hour each. Further examples of high stringency conditions can be found on pages 2.10.1-2.10.16 (see particularly 2.10.8-11) and pages 6.3.1-6 in *Current Protocols in Molecular Biology* (Ausubel, F.M. *et al.*, eds., containing supplements up through Supplement 42, 1998). Examples of high, medium, and low stringency conditions can be found on pages 36 and 37 of WO 98/40404, which are incorporated herein by reference.

The invention further relates to nucleic acids obtainable by screening an appropriate library with a probe having a nucleotide sequence such as that set forth in Figures 44A-44C (SEQ ID NO:46), Figures 46A-46B (SEQ ID NO:48), Figures 94A and 94B (SEQ ID NO:101), Figures 50A-50C (SEQ ID NO:52), Figure 52 (SEQ ID NO:54) or Figures 54A-54C (SEQ ID NO:56), or a probe which is a sufficiently long fragment of any of the above; and isolating the nucleic acid. Such probes generally can comprise at least 15 nucleotides. Nucleic acids obtainable by such screenings may include RNAs, cDNAs and genomic DNA, for example, encoding FATPs of the FATP family described herein.

Further uses for the nucleic acid molecules of the invention, whether encoding a full-length FATP or whether comprising a contiguous portion of a nucleic acid molecule such as one given in SEQ ID NO:46, 48, 101, 52, 54, or 56, include use as markers for tissues in which the corresponding protein is preferentially expressed (to identify constitutively expressed proteins or proteins produced at a particular stage of tissue differentiation or stage of development of a disease state); as molecular weight markers on southern gels; as chromosome markers or tags (when labeled, for example with biotin, a radioactive label or a fluorescent label) to identify chromosomes or to map related gene positions; to compare with endogenous DNA sequences in a mammal to identify potential genetic disorders; as probes to hybridize and thus identify, related DNA sequences; as a source of information to derive PCR primers for genetic fingerprinting; as a probe to "subtract-out" known sequences in the process of

discovering other novel nucleic acid molecules; for selecting and making oligomers for attachment to a "gene chip" or other support, to be used, for example, for examination of expression patterns; to raise anti-protein antibodies using DNA immunization techniques; and as an antigen to raise anti-DNA antibodies or to elicit another immune response.

Further methods to obtain nucleic acids encoding FATPs of the FATP family include PCR and variations thereof (e.g., "RACE" PCR and semi-specific PCR methods). Portions of the nucleic acids having a nucleotide sequence set forth in Figures 44A-44C (SEQ ID NO:46), Figures 46A-46B (SEQ ID NO:48), Figures 94A and 94B (SEQ ID NO:101), Figures 50A-50C (SEQ ID NO:52), Figure 52 (SEQ ID NO:54) or Figures 54A-54C (SEQ ID NO:56), (especially "flanking sequences" on either side of a coding region) can be used as primers in methods using the polymerase chain reaction, to produce DNA from an appropriate template nucleic acid.

Once a fragment of the FATP gene is generated by PCR, it can be sequenced, and the sequence of the product can be compared to other DNA sequences, for example, by using the BLAST Network Service at the National Center for Biotechnology Information. The boundaries of the open reading frame can then be identified using semi-specific PCR or other suitable methods such as library screening. Once the 5' initiator methionine codon and the 3' stop codon have been identified, a PCR product encoding the full-length gene can be generated using genomic DNA as a template, with primers complementary to the extreme 5' and 3' ends of the gene or to their flanking sequences. The full-length genes can then be cloned into expression vectors for the production of functional proteins.

The invention also relates to isolated proteins or polypeptides such as those encoded by nucleic acids of the present invention. Isolated proteins can be purified from a natural source or can be made recombinantly. Proteins or polypeptides referred to herein as "isolated" are proteins or polypeptides that exist in a state different from the state in which they exist in cells in which they are normally expressed in an organism, and include proteins or polypeptides obtained by methods described herein, similar

state in which they exist in cells in which they are normally expressed in an organism, and include proteins or polypeptides obtained by methods described herein, similar methods or other suitable methods, and also include essentially pure proteins or polypeptides, proteins or polypeptides produced by chemical synthesis or by combinations of biological and chemical methods, and recombinant proteins or polypeptides which are isolated. Thus, the term "isolated" as used herein, indicates that the polypeptide in question exists in a physical milieu distinct from that in which it occurs in nature. Thus, "isolated" includes existing in membrane fragments and vesicles membrane fractions, liposomes, lipid bilayers and other artificial membrane systems. An isolated FATP may be substantially isolated with respect to the complex cellular milieu in which it naturally occurs, and may even be purified essentially to homogeneity, for example as determined by PAGE or column chromatography (for example, HPLC), but may also have further cofactors or molecular stabilizers, such as detergents, added to the purified protein to enhance activity. In one embodiment, proteins or polypeptides are isolated to a state at least about 75% pure; more preferably at least about 85% pure, and still more preferably at least about 95% pure, as determined by Coomassie blue staining of proteins on SDS-polyacrylamide gels. Proteins or polypeptides referred to herein as "recombinant" are proteins or polypeptides produced by the expression of recombinant nucleic acids.

In a preferred embodiment, an isolated polypeptide comprising a FATP, a functional portion thereof, or a functional equivalent of the FATP, has at least one function characteristic of a FATP, for example, transport activity, binding function (e.g., a domain which binds to AMP), or antigenic function (e.g., binding of antibodies that also bind to a naturally-occurring FATP, as that function is found in an antigenic determinant). Functional equivalents can have activities that are quantitatively similar to, greater than, or less than, the reference protein. These proteins include, for example, naturally occurring FATPs that can be purified from tissues in which they are produced (including polymorphic or allelic variants), variants (e.g., mutants) of those proteins and/or portions thereof. Such variants include mutants differing by the addition,

deletion or substitution of one or more amino acid residues, or modified polypeptides in which one or more residues are modified, and mutants comprising one or more modified residues. Portions or fragments of a FATP can range in size from four amino acid residues to the entire amino acid sequence minus one amino acid.

The isolated proteins of the invention preferably include mammalian fatty acid transport proteins of the FATP family of homologous proteins. In one embodiment, the extent of amino acid sequence similarity between a polypeptide having one of the amino acid sequences shown in Figure 45 (SEQ ID NO:47), Figure 47 (SEQ ID NO:49), Figures 94A and 94B (SEQ ID NO:102), Figure 51 (SEQ ID NO:53), Figure 53 (SEQ ID NO:55), or Figure 55 (SEQ ID NO:57), and the respective functional equivalents of these polypeptides is at least about 88%. In other embodiments, the degree of amino acid sequence similarity between a FATP and its respective functional equivalent is at least about 91%, at least about 94%, or at least about 97%.

The polypeptides of the invention also include those FATPs encoded by polynucleotides which are orthologous to those polynucleotides, the sequences of which are described herein in whole or in part. FATPs which are orthologs to those described herein by amino acid sequence, in whole or in part, are, for example fatty acid transport proteins 1-6 of dog, rat chimpanzee, monkey, rabbit, guinea pig, baboon and pig, and are also embodiments of the invention.

To determine the percent identity or similarity of two amino acid sequences or of two nucleic acid sequences, the sequences are aligned for optimal comparison purposes (e.g., gaps can be introduced in one or both of a first and a second amino acid or nucleic acid sequence for optimal alignment, and non-homologous (dissimilar) sequences can be disregarded for comparison purposes). In a preferred embodiment, the length of a reference sequence aligned for comparison purposes is at least 30%, preferably at least 40%, more preferably at least 50%, even more preferably at least 60%, and even more preferably at least 70%, 80%, or 90% of the length of the reference sequence. The amino acid residues or nucleotides at corresponding amino acid positions or nucleotide positions are then compared. When a position in the first

sequence is occupied by the same amino acid residue or nucleotide as the corresponding position in the second sequence, then the molecules are identical at that position (as used herein, amino acid or nucleic acid "identity" is equivalent to amino acid or nucleic acid "similarity"). The percent identity between the two sequences is a function of the number of identical positions shared by the sequences, taking into account the number of gaps, and the length of each gap, which need to be introduced for optimal alignment of the two sequences.

The invention also encompasses polypeptides having a lower degree of identity but having sufficient similarity so as to perform one or more of the same functions performed by the polypeptides described herein by amino acid sequence. Similarity for a polypeptide is determined by conserved amino acid substitution. Such substitutions are those that substitute a given amino acid in a polypeptide by another amino acid of like characteristics. Conservative substitutions are likely to be phenotypically silent. Typically seen as conservative substitutions are the replacements, one for another, among the aliphatic amino acids Ala, Val, Leu, and Ile; interchange of the hydroxyl residues Ser and Thr, exchange of the acidic residues Asp and Glu, substitution between the amide residues Asn and Gln, exchange of the basic residues Lys and Arg and replacements among the aromatic residues Phe, Tyr. Guidance concerning which amino acid changes are likely to be phenotypically silent is found in Bowie *et al., Science* *247*:1306-1310 (1990).

TABLE 1. Conservative Amino Acid Substitutions

Aromatic		Phenylalanine	
		Tryptophan	
		Tyrosine	
Hydrophobic		Leucine	
		Isoleucine	
		Valine	
Polar		Glutamine	
		Asparagine	
Basic		Arginine	
		Lysine	
		Histidine	
Acidic		Aspartic Acid	
		Glutamic Acid	
Small		Alanine	
		Serine	
		Threonine	
		Methionine	
		Glycine	

The comparison of sequences and determination of percent identity and similarity between two sequences can be accomplished using a mathematical algorithm. (*Computational Molecular Biology*, Lesk, A.M.,ed., Oxford University Press, New York, 1988; *Biocomputing: Informatics and Genome Projects,* Smith, D.W., ed., Academic Press, New York, 1993; *Computer Analysis of Sequence Data, Part 1,* Griffin, A.M., and Griffin, H.G., eds., Humana Press, New Jersey, 1994; *Sequence*

Analysis in Molecular Biology, von Heinje, G., Academic Press, 1987; and *Sequence Analysis Primer*, Gribskov, M. and Devereaux, J., eds., M. Stockton Press, New York, 1991). In a preferred embodiment, the percent identity between two amino acid sequences is determined using the Needleman and Wunsch (*J. Mol. Biol.* (48):444-453 (1970)) algorithm which has been incorporated into the GAP program in the GCG software package (available at http://www.gcg.com), using either a Blossom 62 matrix or a PAM250 matrix, and a gap weight of 16, 14, 12, 10, 8, 6, or 4 and a length weight of 1, 2, 3, 4, 5, or 6. In yet another preferred embodiment, the percent identity between two nucleotide sequences is determined using the GAP program in the GCG software package (Devereux, J., *et al.*, *Nucleic Acids Res. 12(1)*:387 (1984)) (available at http://www.gcg.com), using a NWSgapdna.CMP matrix and a gap weight of 40, 50, 60, 70, or 80 and a length weight of 1, 2, 3, 4, 5, or 6. In another embodiment, the percent identity between two amino acid or nucleotide sequences is determined using the algorithm of E. Meyers and W. Miller (*CABIOS, 4*:11-17 (1989)) which has been incorporated into the ALIGN program (version 2.0), using a PAM120 weight residue table, a gap length penalty of 12 and a gap penalty of 4.

The nucleic acid and protein sequences of the present invention can further be used as a "query sequence" to perform a search against databases to, for example, identify other family members or related sequences. Such searches can be performed using the NBLAST and XBLAST programs (version 2.0) of Altschul, *et al.* (*J. Mol. Biol. 215*:403-10 (1990)). BLAST nucleotide searches can be performed with the NBLAST program, score = 100, word length = 12 to obtain nucleotide sequences homologous to (with calculatably significant similarity to) the nucleic acid molecules of the invention. BLAST protein searches can be performed with the XBLAST program, score = 50, word length = 3 to obtain amino acid sequences homologous to the proteins of the invention. To obtain gapped alignments for comparison purposes, Gapped BLAST can be utilized as described in Altschul *et al., (Nucleic Acids Res. 25*(17):3389-3402 (1997)). When utilizing BLAST and gapped BLAST programs, the default

parameters of the respective programs (e.g., XBLAST and NBLAST) can be used. See http://www.ncbi.nlm.nih.gov.

Similarity for nucleotide and amino acid sequences can be defined in terms of the parameters set by the Advanced Blast search available from NCBI (the National Center for Biotechnology Information; see, for Advanced BLAST page, www.ncbi.nlm.nih.gov/cgi-bin/BLAST/nph-newblast?Jform=1). These default parameters, recommended for a query molecule of length greater than 85 amino acid residues or nucleotides have been set as follows: gap existence cost, 11, per residue gap cost, 1; lambda ratio, 0.85. Further explanation of version 2.0 of BLAST can be found on related website pages and in Altschul, S.F. *et al., Nucleic Acids Res. 25*:3389-3402 (1997).

The invention further relates to fusion proteins, comprising a FATP or functional portion thereof (as described above) as a first moiety, linked to second moiety not occurring in the FATP as found in nature. Thus, the second moiety can be an amino acid, peptide or polypeptide. The first moiety can be in an N-terminal location, C-terminal location or internal to the fusion protein. In one embodiment, the fusion protein comprises a FATP as the first moiety, and a second moiety comprising a linker sequence and an affinity ligand. Fusion proteins can be produced by a variety of methods. For example, a fusion protein can be produced by the insertion of a FATP gene or portion thereof into a suitable expression vector, such as Bluescript SK +/- (Stratagene), pGEX-4T-2 (Pharmacia), pET-24(+) (Novagen), or vectors of similar construction. The resulting construct can be introduced into a suitable host cell for expression. Upon expression, fusion protein can be purified from cells by means of a suitable affinity matrix (See e.g., *Current Protocols in Molecular Biology*, Ausubel, F.M. *et al.*, eds., Vol. 2, pp. 16.4.1-16.7.8, containing supplements up through Supplement 42, 1998).

The invention also relates to enzymatically produced, synthetically produced, or recombinantly produced portions of a fatty acid transport protein. Portions of a FATP can be made which have full or partial function on their own, or which when mixed

together (though fully, partially, or nonfunctional alone), spontaneously assemble with one or more other polypeptides to reconstitute a functional protein having at least one function characteristic of a FATP.

Fragments of a FATP can be produced by direct peptide synthesis, for example those using solid-phase techniques (Roberge, J.Y. *et al., Science 269*:202-204 (1995); Merrifield, J., *J. Am. Chem. Soc. 85*:2149-2154 (1963)). Protein synthesis can be performed using manual techniques or by automation. Automated synthesis can be carried out using, for instance, an Applied Biosystems 431A Peptide Synthesizer (Perkin Elmer). Various fragments of a FATP can be synthesized separately and combined using chemical methods.

One aspect of the invention is a peptide or polypeptide having the amino acid sequence of a portion of a fatty acid transport protein which is hydrophilic rather than hydrophobic, and ordinarily can be detected as facing the outside of the cell membrane. Such a peptide or polypeptide can be thought of as being an extracellular domain of the FATP, or a mimetic of said extracellular domain. It is known, for example, that a portion of human FATP4 that includes a highly conserved motif is involved in AMP-CoA binding function (Stuhlsatz-Krouper, S.M. *et al., J. Biol. Chem. 44*:28642-28650 (1998)).

The term "mimetic" as used herein, refers to a molecule, the structure of which is developed from knowledge of the structure of the FATP of interest, or one or more portions thereof, and, as such, is able to effect some or all of the functions of a FATP.

Portions of an FATP can be prepared by enzymatic cleavage of the isolated protein, or can be made by chemical synthesis methods. Portions of a FATP can also be made by recombinant DNA methods in which restriction fragments, or fragments that may have undergone further enzymatic processing, or synthetically made DNAs are joined together to construct an altered FATP gene. The gene can be made such that it encodes one or more desired portions of a FATP. These portions of FATP can be entirely homologous to a known FATP, or can be altered in amino acid sequence relative to naturally occurring FATPs to enhance or introduce desired properties such as

solubility, stability, or affinity to a ligand. A further feature of the gene can be a sequence encoding an N-terminal signal peptide directed to the plasma membrane.

An extracellular domain can be determined by a hydrophobicity plot, such as those shown in Figures 28A, 29A, and 35A, or by a hydrophilicity plot such as those shown in Figures 28C, 29C, 35C, 91, 92 and 93. A polypeptide or peptide comprising all or a portion of a FATP extracellular domain can be used in a pharmaceutical composition. When administered to a mammal by an appropriate route, the polypeptide or peptide can bind to fatty acids and compete with the native FATPs in the membrane of cells, thereby making fewer fatty acid molecules available as substrates for transport into cells, and reducing the amount of fatty acids taken up by, for example, the heart, in the case of FATP6.

Another aspect of the invention relates to a method of producing a fatty acid transport protein, variants or portions thereof, and to expression systems and host cells containing a vector appropriate for expression of a fatty acid transport protein.

Cells that express a FATP, a variant or a portion thereof, or an ortholog of a FATP described herein by amino acid sequence, can be made and maintained in culture, under conditions suitable for expression, to produce protein in the cells for cell-based assays, or to produce protein for isolation. These cells can be procaryotic or eucaryotic. Examples of procaryotic cells that can be used for expression include *Escherichia coli, Bacillus subtilis* and other bacteria. Examples of eucaryotic cells that can be used for expression include yeasts such as *Saccharomyces cerevisiae, Schizosaccharomyces pombe, Pichia pastoris* and other lower eucaryotic cells, and cells of higher eucaryotes such as those from insects and mammals, such as primary cells and cell lines such as CHO, HeLa, 3T3 and BHK cells, preferably COS cells and human kidney 293 cells, and more preferably Jurkat cells. (See, e.g., Ausubel, F.M. *et al.,* eds. *Current Protocols in Molecular Biology,* Greene Publishing Associates and John Wiley & Sons, Inc., containing Supplements up through Supplement 42, 1998)).

In one embodiment, host cells that produce a recombinant FATP, or a portion thereof, a variant, or an ortholog of a FATP described herein by amino acid sequence,

can be made as follows. A gene encoding a FATP, variant or a portion thereof can be inserted into a nucleic acid vector, e.g., a DNA vector, such as a plasmid, phage, cosmid, phagemid, virus, virus-derived vector (e.g., SV40, vaccinia, adenovirus, fowl pox virus, pseudorabies viruses, retroviruses) or other suitable replicon, which can be present in a single copy or multiple copies, or the gene can be integrated in a host cell chromosome. A suitable replicon or integrated gene can contain all or part of the coding sequence for a FATP or variant, operably linked to one or more expression control regions whereby the coding sequence is under the control of transcription signals and linked to appropriate translation signals to permit translation. The vector can be introduced into cells by a method appropriate to the type of host cells (e.g., transfection, electroporation, infection). For expression from the FATP gene, the host cells can be maintained under appropriate conditions (e.g., in the presence of inducer, normal growth conditions, etc.). Proteins or polypeptides thus produced can be recovered (e.g., from the cells, as in a membrane fraction, from the periplasmic space of bacteria, from culture medium) using suitable techniques. Appropriate membrane targeting signals may be incorporated into the expressed polypeptide. These signals may be endogenous to the polypeptide or they may be heterologous signals.

Polypeptides of the invention can be recovered and purified from cell cultures (or from their primary cell source) by well-known methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and high performance liquid chromatography. Known methods for refolding protein can be used to regenerate active conformation if the polypeptide is denatured during isolation or purification.

In a further aspect of the invention are methods for assessing the transport function of any of the fatty acid transport proteins or polypeptides described herein, including orthologs, and in variations of these, methods for identifying an inhibitor (or an enhancer) of such function and methods for assessing the transport function in the presence of a candidate inhibitor or a known inhibitor.

A variety of systems comprising living cells can be used for these methods. Cells to be used in fatty acid transport assays, and further in methods for identifying an inhibitor or enhancer of this function, express one or more FATPs. See Examples 3, 6, 9, 12 and 14 for data on tissue distribution of expression of FATPs, and Examples 10 and 11 describing recombinant cells expressing FATP. Cells for use in cell-based assays described herein can be drawn from a variety of sources, such as isolated primary cells of various organs and tissues wherein one or more FATPs are naturally expressed. In some cases, the cells can be from adult organs, and in some cases, from embryonic or fetal organs, such as heart, lung, liver, intestine, skeletal muscle, kidney and the like. Cells for this purpose can also include cells cultured as fragments of organs or in conditions simulating the cell type and/or tissue organization of organs, in which artificial materials may be used as substrates for cell growth. Other types of cells suitable for this purpose include cells of a cell strain or cell line (ordinarily comprising cells considered to be "transformed") transfected to express one or more FATPs.

A further embodiment of the invention is a method for detecting, in a sample of cells, a fatty acid transport protein, a portion or fragment thereof, a fusion protein comprising a FATP or a portion thereof, or an ortholog as described herein, wherein the cells can be, for instance, cells of a tissue, primary culture cells, or cells of a cell line, including cells into which nucleic acid has been introduced. The method comprises adding to the sample an agent that specifically binds to the protein, and detecting the agent specifically bound to the protein. Appropriate washing steps can be added to reduce nonspecific binding to the agent. The agent can be, for example, an antibody, a ligand or a substrate mimic. The agent can have incorporated into it, or have bound to it, covalently or by high affinity non-covalent interactions, for instance, a label that facilitates detection of the agent to which it is bound, wherein the label can be, but is not limited to, a phosphorescent label, a fluorescent label, a biotin or avidin label, or a radioactive label. The means of detection of a fatty acid transport protein can vary, as appropriate to the agent and label used. For example, for an antibody that binds to the fatty acid transport protein, the means of detection may call for binding a second

antibody, which has been conjugated to an enzyme, to the antibody which binds the fatty acid transport protein, and detecting the presence of the second antibody by means of the enzymatic activity of the conjugated enzyme.

Similar principles can also be applied to a cell lysate or a more purified preparation of proteins from cells that may comprise a fatty acid transport protein of interest, for example in the methods of immunoprecipitation, immunoblotting, immunoaffinity methods, that in addition to detection of the particular FATP, can also be used in purification steps, and qualitative and quantitative immunoassays. See, for instance, chapters 11 through 14 in *Antibodies: A Laboratory Manual*, E. Harlow and D. Lane, eds., Cold Spring Harbor Laboratory, 1988.

Isolated fatty acid transport protein or, an antigenically similar portion thereof, especially a portion that is soluble, can be used in a method to select and identify molecules which bind specifically to the FATP. Fusion proteins comprising all of, or a portion of, the fatty acid transport protein linked to a second moiety not occurring in the FATP as found in nature, can be prepared for use in another embodiment of the method. Suitable fusion proteins for this purpose include those in which the second moiety comprises an affinity ligand (e.g., an enzyme, antigen, epitope). FATP fusion proteins can be produced by the insertion of a gene encoding the FATP or a variant thereof, or a suitable portion of such gene into a suitable expression vector, which encodes an affinity ligand (e.g., pGEX-4T-2 and pET-15b, encoding glutathione S-transferase and His-Tag affinity ligands, respectively). The expression vector can be introduced into a suitable host cell for expression. Host cells are lysed and the lysate, containing fusion protein, can be bound to a suitable affinity matrix by contacting the lysate with an affinity matrix.

In one embodiment, the fusion protein can be immobilized on a suitable affinity matrix under conditions sufficient to bind the affinity ligand portion of the fusion protein to the matrix, and is contacted with one or more candidate binding agents (e.g., a mixture of peptides) to be tested, under conditions suitable for binding of the binding agents to the FATP portion of the bound fusion protein. Next, the affinity matrix with

bound fusion protein can be washed with a suitable wash buffer to remove unbound candidate binding agents and non-specifically bound candidate binding agents. Those agents which remain bound can be released by contacting the affinity matrix with fusion protein bound thereto with a suitable elution buffer. Wash buffer can be formulated to permit binding of the fusion protein to the affinity matrix, without significantly disrupting binding of specifically bound binding agents. In this aspect, elution buffer can be formulated to permit retention of the fusion protein by the affinity matrix, but can be formulated to interfere with binding of the candidate binding agents to the target portion of the fusion protein. For example, a change in the ionic strength or pH of the elution buffer can lead to release of specifically bound agent, or the elution buffer can comprise a release component or components designed to disrupt binding of specifically bound agent to the target portion of the fusion protein.

Immobilization can be performed prior to, simultaneous with, or after, contacting the fusion protein with candidate binding agent, as appropriate. Various permutations of the method are possible, depending upon factors such as the candidate molecules tested, the affinity matrix-ligand pair selected, and elution buffer formulation. For example, after the wash step, fusion protein with binding agent molecules bound thereto can be eluted from the affinity matrix with a suitable elution buffer (a matrix elution buffer, such as glutathione for a GST fusion). Where the fusion protein comprises a cleavable linker, such as a thrombin cleavage site, cleavage from the affinity ligand can release a portion of the fusion with the candidate agent bound thereto. Bound agent molecules can then be released from the fusion protein or its cleavage product by an appropriate method, such as extraction.

One or more candidate binding agents can be tested simultaneously. Where a mixture of candidate binding agents is tested, those found to bind by the foregoing processes can be separated (as appropriate) and identified by suitable methods (e.g., PCR, sequencing, chromatography). Large libraries of candidate binding agents (e.g., peptides, RNA oligonucleotides) produced by combinatorial chemical synthesis or by other methods can be tested (see e.g., Ohlmeyer, M.H.J. *et al.*, *Proc. Natl. Acad. Sci.*

USA 90:10922-10926 (1993) and DeWitt, S.H. *et al.*, *Proc. Natl. Acad. Sci. USA 90*:6909-6913 (1993), relating to tagged compounds; see also Rutter, W.J. *et al.* U.S. Patent No. 5,010,175; Huebner, V.D. *et al.*, U.S. Patent No. 5,182,366; and Geysen, H.M., U.S. Patent No. 4,833,092). Random sequence RNA libraries (see Ellington, A.D. *et al., Nature 346*:818-822 (1990); Bock, L.C. *et al., Nature 355*:584-566 (1992); and Szostak, J.W., *Trends in Biochem. Sci. 17*:89-93 (March, 1992)) can also be screened according to the present method to select RNA molecules which bind to a target FATP or FATP fusion protein. Where binding agents selected from a combinatorial library by the present method carry unique tags, identification of individual biomolecules by chromatographic methods is possible. Where binding agents do not carry tags, chromatographic separation, followed by mass spectrometry to ascertain structure, can be used to identify binding agents selected by the method, for example.

The invention also comprises a method for identifying an agent which inhibits interaction between a fatty acid transport protein (e.g., one comprising the amino acid sequence in SEQ ID NO:47, SEQ ID NO:49, SEQ ID NO:51, SEQ ID NO:53, SEQ ID NO:102, or SEQ ID NO:57), and a ligand of said protein. The FATP can be one described by amino acid sequence herein, a portion or fragment thereof, a variant thereof, or an ortholog thereof, or a FATP fusion protein. Here, a ligand can be, for instance, a substrate, or a substrate mimic, an antibody, or a compound, such as a peptide, that binds with specificity to a site on the protein. The method comprises combining, not limited to a particular order, the fatty acid protein, the ligand of the protein, and a candidate agent to be assessed for its ability to inhibit interaction between the protein and the ligand, under conditions appropriate for interaction between the protein and the ligand (e.g., pH, salt, temperature conditions conducive to appropriate conformation and molecular interactions); determining the extent to which the protein and ligand interact; and comparing (1) the extent of protein-ligand interaction in the presence of candidate agent with (2) the extent of protein-ligand interaction in the

absence of candidate agent, wherein if (1) is less than (2), then the candidate agent is one which inhibits interaction between the protein and the ligand.

The method can be facilitated, for example, by using an experimental system which employs a solid support (column chromatography matrix, wall of a plate, microtiter wells, column pore glass, pins to be submerged in a solution, beads, etc.) to which the protein can be attached. Accordingly, in one embodiment, the protein can be fixed to a solid phase directly or indirectly, by a linker. The candidate agent to be tested is added under conditions conducive for interaction and binding to the protein. The ligand is added to the solid phase system under conditions appropriate for binding. Excess ligand is removed, as by a series of washes done under conditions that do not disrupt protein-ligand interactions. Detection of bound ligand can be facilitated by using a ligand that carries a label (e.g., fluorescent, chemiluminescent, radioactive). In a control experiment, protein and ligand are allowed to interact in the absence of any candidate agent, under conditions otherwise identical to those used for the "test" conditions where candidate inhibiting agent is present, and any washes used in the test conditions are also used in the control. The extent to which ligand binds to the protein in the presence of candidate agent is compared to the extent to which ligand binds to the protein in the absence of the candidate agent. If the extent to which interaction of the protein and the ligand occurs is less in the presence of the candidate agent than in the absence of the candidate agent, the candidate agent is an agent which inhibits interaction between the protein and the ligand of the protein.

In a further embodiment, an inhibitor (or an enhancer) of a fatty acid transport protein can be identified. The method comprises steps which are, or are variations of the following: contacting the cells with fatty acid, wherein the fatty acid can be labeled for convenience of detection; contacting a first aliquot of the cells with an agent being tested as an inhibitor (or enhancer) of fatty acid uptake while maintaining a second aliquot of cells under the same conditions but without contact with the agent; and measuring (e.g., quantitating) fatty acid in the first and second aliquots of cells; wherein a lesser quantity of fatty acid in the first aliquot compared to that in the second aliquot

is indicative that the agent is an inhibitor of fatty acid uptake by a fatty acid transport protein. A greater quantity of fatty acid in the first aliquot compared to that in the second aliquot is indicative that the agent is an enhancer of fatty acid uptake by a fatty acid transport protein.

A particular embodiment of identifying an inhibitor or enhancer of fatty acid transport function employs the above steps, but also employs additional steps preceding those given above: introducing into cells of a cell strain or cell line ("host cells" for the intended introduction of, or after the introduction of, a vector) a vector comprising a fatty acid transport protein gene, wherein expression of the gene can be regulatable or constitutive, and providing conditions to the host cells under which expression of the gene can occur.

The terms "contacting" and "combining" as used herein in the context of bringing molecules into close proximity to each other, can be accomplished by conventional means. For example, when referring to molecules that are soluble, contacting is achieved by adding the molecules together in a solution. "Contacting" can also be adding an agent to a test system, such as a vessel containing cells in tissue culture.

The term "inhibitor" or "antagonist", as used herein, refers to an agent which blocks, diminishes, inhibits, hinders, limits, decreases, reduces, restricts or interferes with fatty acid transport into the cytoplasm of a cell, or alternatively and additionally, prevents or impedes the cellular effects associated with fatty acid transport. The term "enhancer" or "agonist", as used herein, refers to an agent which augments, enhances, or increases fatty acid transport into the cytoplasm of a cell. An antagonist will decrease fatty acid concentration, fatty acid metabolism and byproduct levels in the cell, leading to phenotypic and molecular changes.

In order to produce a "host cell" type suitable for fatty acid uptake assays and for assays derived therefrom for identifying inhibitors or enhancers thereof, a nucleic acid vector can be constructed to comprise a gene encoding a fatty acid transport protein, for example, human FATP1, FATP2, FATP3, FATP4, FATP5, FATP6, a mutant or variant

thereof, an ortholog of the human proteins, such as mouse orthologs or orthologs found in other mammals, or a FATP family protein of origin in an organism other than a mammal. The gene of the vector can be regulatable, such as by the placement of the gene under the control of an inducible or repressible promoter in the vector (e.g., inducible or repressible by a change in growth conditions of the host cell harboring the vector, such as addition of inducer, binding or functional removal of repressor from the cell millieu, or change in temperature) such that expression of the FATP gene can be turned on or initiated by causing a change in growth conditions, thereby causing the protein encoded by the gene to be produced, in host cells comprising the vector, as a plasma membrane protein. Alternatively, the FATP gene can be constitutively expressed.

A vector comprising an FATP gene, such as a vector described herein, can be introduced into host cells by a means appropriate to the vector and to the host cell type. For example, commonly used methods such as electroporation, transfection, for instance, transfection using $CaCl_2$, and transduction (as for a virus or bacteriophage) can be used. Host cells can be, for example, mammalian cells such as primary culture cells or cells of cell lines such as COS cells, 293 cells or Jurkat cells. Host cells can also be, in some cases, cells derived from insects, cells of insect cell lines, bacterial cells, such as *E. coli*, or yeast cells, such as *S. cerevisiae.* It is preferred that the fatty acid transport protein whose function is to be assessed, with or without a candidate inhibitor or enhancer, be produced in host cells whose ancestor cells originated in a species related to the species of origin of the FATP gene encoding the fatty acid transport protein. For example, it is preferable that tests of function or of inhibition or enhancement of a mammalian FATP be carried out in host mammalian cells producing the FATP, rather than bacterial cells or yeast cells.

Host cells comprising a vector comprising a regulatable FATP gene can be treated so as to allow expression of the FATP gene and production of the encoded protein (e.g., by contacting the cells with an inducer compound that effects transcription from an inducible promoter operably linked to the FATP gene).

The test agent (e.g., an agonist or antagonist) is added to the cells to be used in a fatty acid transport assay, in the presence or absence of test agent, under conditions suitable for production and/or maintenance of the expressed FATP in a conformation appropriate for association of the FATP with test agent and substrate. For example, conditions under which an agent is assessed, such as media and temperature requirements, can, initially, be similar to those necessary for transport of typical fatty acid substrates across the plasma membrane. One of ordinary skill in the art will know how to vary experimental conditions depending upon the biochemical nature of the test agent. The test agent can be added to the cells in the presence of fatty acid, or in the absence of fatty acid substrate, with the fatty acid substrate being added following the addition of the test agent. The concentration at which the test agent can be evaluated can be varied, as appropriate, to test for an increased effect with increasing concentrations.

Test agents to be assessed for their effects on fatty acid transport can be any chemical (element, molecule, compound), made synthetically, made by recombinant techniques or isolated from a natural source. For example, test agents can be peptides, polypeptides, peptoids, sugars, hormones, or nucleic acid molecules, such as antisense nucleic acid molecules. In addition, test agents can be small molecules or molecules of greater complexity made by combinatorial chemistry, for example, and compiled into libraries. These libraries can comprise, for example, alcohols, alkyl halides, amines, amides, esters, aldehydes, ethers and other classes of organic compounds. Test agents can also be natural or genetically engineered products isolated from lysates of cells, bacterial, animal or plant, or can be the cell lysates themselves. Presentation of test compounds to the test system can be in either an isolated form or as mixtures of compounds, especially in initial screening steps.

Thus, the invention relates to a method for identifying agents which alter fatty acid transport, the method comprising providing the test agent to the cell (wherein "cell" includes the plural, and can include cells of a cell strain, cell line or culture of primary cells or organ culture, for example), under conditions suitable for binding to its target,

whether to the FATP itself or to another target on or in the cell, wherein the transformed cell comprises a FATP.

In greater detail, to test one or more agents or compounds (e.g., a mixture of compounds can conveniently be screened initially) for inhibition of the transport function of a fatty acid transport protein, the agent(s) can be contacted with the cells. The cells can be contacted with a labeled fatty acid. The fatty acid can be, for example, a known substrate of the fatty acid transport protein such as oleate or palmitate. The fatty acid can itself be labeled with a radioactive isotope, (e.g., ^{3}H or ^{14}C) or can have a radioactively labeled adduct attached. In other variations, the fatty acid can have chemically attached to it a fluorescent label, or a substrate for an enzyme occurring within the cells, wherein the substrate yields a detectable product, such as a highly colored or fluorescent product. Addition of candidate inhibitors and labeled substrate to the cells comprising fatty acid transport protein can be in either order or can be simultaneous.

A second aliquot of cells, which can be called "control" cells (a "first" aliquot of cells can be called "test" cells), is treated, if necessary (as in the case of transformed "host"cells), so as to allow expression of the FATP gene, and is contacted with the labeled substrate of the fatty acid transport protein. The second aliquot of cells is not contacted with one or more agents to be tested for inhibition of the transport function of the protein produced in the cells, but is otherwise kept under the same culture conditions as the first aliquot of cells.

In a further step of a method to identify inhibitors of a fatty acid transport protein, the labeled fatty acid is measured in the first and second aliquots of cells. A preliminary step of this measurement process can be to separate the external medium from the cells so as to be able to distinguish the labeled fatty acid external to the cells from that which has been transported inside the cells. This can be accomplished, for instance, by removing the cells from their growth container, centrifuging the cell suspension, removing the supernatant and performing one or more wash steps to extensively dilute the remaining medium which may contain labeled fatty acid.

Detection of the labeled fatty acid can be by a means appropriate to the label used. For example, for a radioactive label, detection can be by scintillation counting of appropriately prepared samples of cells (e.g., lysates or protein extracts); for a fluorescent label, by measuring fluorescence in the cells by appropriate instrumentation.

If a compound tested as a candidate inhibitor of transport function causes the test cells to have less labeled fatty acid detected in the cells than that detected in the control cells, then the compound is an inhibitor of the fatty acid transport protein. Procedures analogous to those above can be devised for identifying enhancers (agonists of FATPs) of fatty acid transport function wherein if the test cells contain more labeled fatty acid than that detected in the control cells, or if the fatty acid is taken up at a higher rate, then the compound being tested can be concluded to be an enhancer of the fatty acid transport protein.

Example 13 describes use of an assay of this type to identify an inhibitor of a FATP. In Example 13, an antisense oligonucleotide which specifically inhibits biosynthesis of mmFATP4 was demonstrated to inhibit fatty acid uptake into mouse enterocytes. Similarly, antisense oligonucleotides directed towards specifically inhibiting the biosynthesis of FATP6 in heart cells, FATP5 in liver cells, FATP3 in lung cells, and FATP2 in colon cells, can be demonstrated as examples of "test agents" that inhibit fatty acid transport.

Another assay to determine whether an agent is an inhibitor (or enhancer) of fatty acid transport employs animals, one or more of which are administered the agent, and one or more of which are maintained under similar conditions, but are not administered the agent. Both groups of animals are given fatty acids (e.g., orally, intravenously, by tube inserted into stomach or intestine), and the fatty acids taken up into a bodily fluid (e.g., serum) or into an organ or tissue of interest are measured from comparable samples taken from each group of animals. The fatty acids may carry a label (e.g., radioactive) to facilitate detection and quantitation of fatty acids taken up into the fluid or tissue being sampled. This type of assay can be used alone or can be used in addition to *in vitro* assays of a candidate inhibitor or enhancer.

An agent determined to be an inhibitor (or enhancer) of FATP function, such as fatty acid binding and/or fatty acid uptake, can be administered to cells in culture, or *in vivo*, to a mammal (e.g. human) to inhibit (or enhance) FATP function. Such an agent may be one that acts directly on the FATP (for example, by binding) or can act on an intermediate in a biosynthetic pathway to produce FATP, such as transcription of the FATP gene, processing of the mRNA, or translation of the mRNA. An example of such an agent is antisense oligonucleotide.

Antisense methods similar to those illustrated in Example 13 can be used to determine the target FATP of a compound or agent that has an inhibitory or enhancing effect on fatty acid uptake. For example, antisense oligonucleotide directed to the inhibition of FATP4 biosynthesis can be added to lung cells or cell lines derived from lung cells. In addition, antisense oligonucleotides directed to the inhibition of other FATPs, except for FATP3, can also be added to the lung cells. The administration of antisense oligonucleotides in this manner ensures that the predominant FATP activity remaining in the cells comes from FATP3. After a period of incubation of the cells with the antisense oligonucleotides sufficient to deplete the plasma membrane of the FATPs whose biosynthesis has been inhibited, a test agent, preferably one that has been shown by some preliminary test to have an inhibitory or enhancing activity on fatty acid transport, can be added to the lung cells. If the test agent is now demonstrated, after treatment of the cells with antisense oligonucleotides, to have an inhibitory or enhancing activity on fatty acid transport in the lung cells, it can be concluded that the target of the test agent is FATP3, or a molecule involved in the biosynthesis or activity of FATP3.

In another type of cell-based assay for uptake of fatty acids, a change of intracellular pH resulting from the uptake of fatty acids can be followed by an indicator fluorophore. The fluorophore can be taken up by the cells in a preincubation step. Fatty acids can be added to the cell medium, and after some period of incubation to allow FATP-mediated uptake of fatty acids, the change in λ_{max} of fluorescence can be measured, as an indicator of a change in intracellular pH, as the λ_{max} of fluorescence of

the fluorophore changes with the pH of its environment, thereby indicating uptake of fatty acids. One such fluorophore is BCECF (2′, 7′-bis(2-carboxyethyl)-5(6)-carboxyfluorescein; Rink, T.J. *et al., J. Cell. Biol. 95*: 189 (1982)).

In assays similar to those described above, a candidate inhibitor or enhancer of fatty acid transport function can be added (or mock-added, for control cultures) to cultures of cells engineered to express a desired FATP to which fatty acid substrate is also added. Inhibition of fatty acid uptake is indicated by a lack of the drop in pH, indicating fatty acid uptake, that is seen in control cells. Enhancement of fatty acid uptake is indicated by a decrease in intracellular pH, as compared to control cells not receiving the candidate enhancer of fatty acid transport function.

Yeast cells can be used in a similar cell-based assay for the uptake of fatty acids mediated by a FATP, and such an assay can be adapted to a screening assay for the identification of agents that inhibit or enhance fatty acid uptake by an FATP. Yeast cells lacking an endogenous FATP activity (mutated, disrupted or deleted for *FAT1*; Faergeman, N.J. *et al., J. Biol. Chem. 272*(13):8531-8538 (1997); Watkins, P.A. *et al., J. Biol. Chem. 273*(29):18210-18219 (1998)) can be engineered to harbor a related gene of the family of FATP-encoding genes, such as a mammalian FATP (e.g., human FATP4).

Examples of expression vectors include pEG (Mitchell, D.A., *et al., Yeast 9*:715-723 (1993)) and pDAD1 and pDAD2, which contain a *GAL1* promoter (Davis, L. I. and Fink, G. R., *Cell 61*:965-978 (1990)). A variety of promoters are suitable for expression. Available yeast vectors offer a choice of promoters. In one embodiment, the inducible *GAL1* promoter is used. In another embodiment, the constitutive *ADH1* promoter (alcohol dehyrodrogenase; Bennetzen, J. L. and Hall, B. D., *J. Biol. Chem. 257*:3026-3031 (1982)) can be used to express an inserted gene on glucose-containing media. An example of a vector suitable for expression of a heterologous FATP gene in yeast is pQB169.

With the introduced FATP gene providing the only fatty acid transport protein function for the yeast cells, it is possible to study effect of the heterologous FATP on

fatty acid transport into the yeast cells in isolation. Assays for the uptake of fatty acids into the yeast cells can be devised that are similar to those described above and/or those assays that have been illustrated in the Examples. Tests for candidate inhibitors or enhancers of the heterologous FATP can be done in cultures of yeast cells, wherein the yeast cells are incubated with fatty acid substrate and an agent to be tested as an inhibitor or enhancer of FATP function. FATP uptake after a period of time can be measured by analyzing the contents of the yeast cells for fatty acid substrate, as compared with control yeast cells incubated with the fatty acid, but not with the test agent. Yeast cells have the additional advantage, over mammalian cells in culture, for example, that yeast cells can be forced to rely upon fatty acids as their only source of carbon, if the growth medium supplied to the yeast cells is formulated to contain no other source of carbon. Thus, the effect of the heterologous FATP on fatty acid uptake and metabolism in the engineered yeast cells can be amplified. An agent that efficiently blocks transport function of the heterologous FATP could result in death of the yeast cells. Thus, in this case, inhibition of function of the heterologous FATP can result in loss of viability. A simple measure of viability is turbidity of the yeast suspension culture, which can be adapted to a high throughput screening assay for effects of various agents to be tested, using microtiter plates or similar devices for small-volume cultures of the engineered yeast cells.

Cell-free assays can also be used to measure the transport of fatty acids across a membrane, and therefor also to assess a test treatment or test agent for its effect on the rate or extent of fatty acid transport. An isolated FATP, for example in the presence of a detergent that preserves the native 3-dimensional structure of the FATP, or partially purified FATP, can be used in an artificial membrane system typically used to preserve the native conformation and activity of membrane proteins. Such systems include liposomes, artificial bilayers of phospholipids, isolated plasma membrane such as cell membrane fragments, cell membrane fractions, or cell membrane vesicles, and other systems in which the FATP can be properly oriented within the membrane to have transport activity. Assays for transport activity can be performed using methods

analogous to those that can be used in cells engineered to predominantly express one FATP whose function is to be measured. A labeled (e.g., radioactively labeled) fatty acid substrate can be incubated with one side of a bilayer or in a suspension of liposomes constructed to integrate a properly oriented FATP. The accumulation of fatty acids with time can be measured, using appropriate means to detect the label (e.g., scintillation counting of medium on each side of the bilayer, or of the contents of liposomes isolated from the surrounding medium). Assays such as these can be adapted to use for the testing of agents which might interact with the FATP to produce an inhibitory or an enhancing effect on the rate or extent of fatty acid transport. That is, the above-described assay can be done in the presence or absence of the agent to be tested, and the results compared.

For examples of isolation of membrane proteins (ADP/ATP carrier and uncoupling protein), reconstitution into phospholipid vesicles, and assays of transport, see Klingenberg, M. *et al., Methods Enzymol. 260*:369-389 (1995). For an example of a membrane protein (phosphate carrier of *Saccharomyces cerevisiae*) that was purified and solubilized from *E. coli* inclusion bodies, see Schroer, A. *et al., J. Biol. Chem. 273*: 14269-14276 (1998). The Glut1 glucose transporter of rat has been expressed in yeast. A crude membrane fraction of the yeast was prepared and reconstituted with soybean phospholipids into liposomes. Glucose transport activity could be measured in the liposomes (Kasahara, T. and Kasahara, M., *J. Biol. Chem. 273*: 29113-29117 (1998)). Similar methods can be applied to the proteins and polypeptides of the invention.

Another embodiment of the invention is a method for inhibiting fatty acid uptake in a mammal (e.g., a human), comprising administering to the mammal a therapeutically effective amount of an inhibitor of the transport function of one or more of the fatty acid transport proteins, thereby decreasing fatty acid uptake by cells comprising the fatty acid protein(s). Where it is desirable to reduce the uptake of fatty acids, for example, in the treatment of chronic obesity or as a part of a program of weight control or hyperlipidemia control in a human, one or more inhibitors of one or more of the fatty acid transport proteins can be administered in an effective dose, and by

an effective route, for example, orally, or by an indwelling device that can deliver doses to the small intestine. The inhibitor can be one identified by methods described herein, or can be one that is, for instance, structurally related to an inhibitor identified by methods described herein (e.g., having chemical adducts to better stabilize or solubilize the inhibitor). The invention further relates to compositions comprising inhibitors of fatty acid uptake in a mammal, which may further comprise pharmaceutical carriers suitable for administration to a subject mammal, such as sterile solubilizing or emulsifying agents.

A further embodiment of the present invention is a method of enhancing or increasing fatty acid uptake, such as enhancing or increasing LCFA uptake in the small intestine (e.g., to treat or prevent a malabsorption syndrome or other wasting condition) or in the liver (e.g., by an enhancer of FATP5 transport activity to treat acute liver failure) or in the kidney (e.g., by an enhancer of FATP2 transport activity to treat kidney failure). In this embodiment, a therapeutically effective amount of an enhancer of the transport function of one or more of the fatty acid transport proteins can be administered to a mammalian subject, with the result that fatty acid uptake in the small intestine is enhanced. In this embodiment, one or more enhancers of one or more of fatty acid transport proteins is administered in an effective dose and by a route (e.g., orally or by a device, such as an indwelling catheter or other device) which can deliver doses to the gut. The enhancer of FATP function (e.g., an enhancer of FATP4 function) can be identified by methods described herein or can be one that is structurally similar to an enhancer identified by methods described herein.

Aerobic reperfusion of ischemic myocardium is a common clinical event which can occur during such treatments as cardiac surgery, angioplasty, and thrombolytic therapy after a myocardial infarction. During reperfusion, a rapid recovery of myocardial energy production is essential for the complete recovery of contractile function. Not only the extent of recovery of myocardial energy metabolism but also the type of energy substrate used by the heart during reperfusion are important determinants of functional recovery. Circulating fatty acid levels increase following acute

myocardial infarction or during cardiac surgery, such that during and following ischemia the heart muscle can be exposed to very high concentrations of fatty acids (Lopaschuk, G.D. and W. C. Stanley, *Science and Medicine* (November/December 1997)). High plasma fatty acid concentrations increase the severity of ischemic damage in a number of experimental models of cardiac ischemia and have been linked to depression of mechanical function during aerobic reperfusion of previously ischemic hearts. Further data show that modifying fatty acid utilization can be beneficial for heart function in ischemia and can be a useful approach for the treatment of angina. See, e.g., Desideri and Celegon, *Am. J. Cardiol. 82(5A)*:50K-53K; Lopaschuk, *Am. J. Cardiol. 82(5A)*:14K-17K. Plasma fatty acid concentrations can be reduced by administering to a human subject or other mammal an effective amount of an inhibitor of a FATP such as FATP2 or FATP4, thereby providing a way of reducing fatty acid utilization by the heart.

In a further embodiment of the invention, a therapeutically effective amount of an inhibitor of hsFATP6 can be administered to a human patient by a suitable route, to reduce the uptake of fatty acids by cardiac muscle. This treatment is desirable in patients who are diagnosed as having, or who are at risk of, abnormal accumulations of fatty acids in the heart or a detrimentally high rate of uptake of fatty acids into the heart, because of ischemic heart disease, or following ischemia or trauma to the heart.

The invention further relates to antibodies that bind to an isolated or recombinant fatty acid transport protein of the FATP family, including portions of antibodies, which can specifically recognize and bind to one or more FATPs. The antibodies and portions thereof of the invention include those which bind to one or more FATPs of mouse or other mammalian species. In a preferred embodiment, the antibodies specifically bind to a naturally occurring FATP of humans. The antibodies can be used in methods to detect or to purify a protein of the present invention or a portion thereof by various methods of immunoaffinity chromatography, to inhibit the function of a protein in a method of therapy, or to selectively inactivate an active site, or to study other aspects of the structure of these proteins, for example.

The antibodies of the present invention can be polyclonal or monoclonal. The term antibody is intended to encompass both polyclonal and monoclonal antibodies. Antibodies of the present invention can be raised against an appropriate immunogen, including proteins or polypeptides of the present invention, such as an isolated or recombinant FATP1, FATP2, FATP3, FATP4, FATP5, FATP6, mtFATP, ceFATPa, ceFATPb, scFATP or portions thereof, or synthetic molecules, such as synthetic peptides (e.g., conjugated to a suitable carrier). Preferred embodiments are antibodies that bind to any of the following: hsFATP1, hsFATP2, hsFATP3, hsFATP4, hsFATP5 or hsFATP6. The immunogen can be a polypeptide comprising a portion of a FATP and having at least one function of a fatty acid transport protein, as described herein.

The term antibody is also intended to encompass single chain antibodies, chimeric, humanized or primatized (CDR-grafted) antibodies and the like, as well as chimeric or CDR-grafted single chain antibodies, comprising portions from more than one species. For example, the chimeric antibodies can comprise portions of proteins derived from two different species, joined together chemically by conventional techniques or prepared as a single contiguous protein using genetic engineering techniques (e.g., DNA encoding the protein portions of the chimeric antibody can be expressed to produce a contiguous protein chain. See, e.g., Cabilly et al., U.S. Patent No. 4,816,567; Cabilly *et al.*, European Patent No. 0,125,023 B1; Boss *et al.*, U.S. Patent No. 4,816,397; Boss *et al.*, European Patent No. 0,120,694 B1; Neuberger, M.S. *et al.*, WO 86/01533; Neuberger, M.S. *et al.*, European Patent No. 0,194,276 B1; Winter, U.S. Patent No. 5,225,539; Winter, European Patent No. 0,239,400 B1; Queen *et al.*, U.S. Patent No. 5,585,089; and Queen *et al.*, European Patent No. EP 0 451 216 B1. See also, Newman, R. *et al., BioTechnology, 10:*1455-1460 (1992), regarding primatized antibody, and Ladner *et al.*, U.S. Patent No. 4,946,778 and Bird, R.E. *et al., Science, 242:*423-426 (1988) regarding single chain antibodies.)

Whole antibodies and biologically functional fragments thereof are also encompassed by the term antibody. Biologically functional antibody fragments which can be used include those fragments sufficient for binding of the antibody fragment to a

FATP to occur, such as Fv, Fab, Fab' and F(ab')$_2$ fragments. Such fragments can be produced by enzymatic cleavage or by recombinant techniques. For instance, papain or pepsin cleavage can generate Fab or F(ab')$_2$ fragments, respectively. Antibodies can also be produced in a variety of truncated forms using antibody genes in which one or more stop codons have been introduced upstream of the natural stop site. For example, a chimeric gene encoding a F(ab')$_2$ heavy chain portion can be designed to include DNA sequences encoding the CH_1 domain and hinge region of the heavy chain.

Preparation of immunizing antigen (whole cells comprising FATP on the cell surface or purified FATP), and polyclonal and monoclonal antibody production can be performed using any suitable technique. A variety of methods have been described (See e.g., Kohler *et al., Nature, 256:* 495-497 (1975) and *Eur. J. Immunol. 6:* 511-519 (1976); Milstein *et al., Nature 266: 550-552 (1977);* Koprowski *et al.,* U.S. Patent No. 4,172,124; Harlow, E. and D. Lane, 1988, *Antibodies: A Laboratory Manual*, (Cold Spring Harbor Laboratory: Cold Spring Harbor, NY); Chapter 11 In *Current Protocols In Molecular Biology*, Vol. 2 (containing supplements up through Supplement 42, 1998), Ausubel, F.M. *et al.*, eds., (John Wiley & Sons: New York, NY)). Generally, a hybridoma can be produced by fusing a suitable immortal cell line (e.g., a myeloma cell line such as SP2/0) with antibody producing cells. The antibody producing cells, preferably those obtained from the spleen or lymph nodes, can be obtained from animals immunized with the antigen of interest. Immunization of animals can be by introduction of whole cells comprising fatty acid transport protein on the cell surface. The fused cells (hybridomas) can be isolated using selective culture conditions, and cloned by limiting dilution. Cells which produce antibodies with the desired specificity can be selected by a suitable assay (e.g., ELISA).

Other suitable methods of producing or isolating antibodies (including human antibodies) of the requisite specificity can used, including, for example, methods which select recombinant antibody from a library (e.g., Hoogenboom *et al.*, WO 93/06213; Hoogenboom *et al.*, U.S. Patent No. 5,565,332; WO 94/13804, published June 23, 1994; and Dower, W.J. *et al.*, U.S. Patent No. 5,427,908), or which rely upon

immunization of transgenic animals (e.g., mice) capable of producing a full repertoire of human antibodies (see e.g., Jakobovits *et al., Proc. Natl. Acad. Sci. USA, 90*: 2551-2555 (1993); Jakobovits *et al., Nature, 362*:255-258 (1993); Lonberg *et al.*, U.S. Patent No. 5,569,825; Lonberg *et al.*, U.S. Patent No. 5,545,806; Surani *et al.*, U.S. Patent No. 5,545,807; and Kucherlapati, R. *et al.*, European Patent No. EP 0 463 151 B1).

Another aspect of the invention is a method for directing an agent to cardiac muscle. The differential expression of FATP6 in cardiac muscle but not in other tissue types allows for the specific targeting of drugs, diagnostic agents, tagging labels, histological stains or other substances specifically to cardiac muscle. A targeting vehicle can be used for the delivery of such a substance. Targeting vehicles which bind specifically to FATP6 can be linked to a substance to be delivered to the cells of cardiac muscle. The linkage can be, for instance, via one or more covalent bonds, or by high affinity non-covalent bonds. A targeting vehicle can be an antibody, for instance, or other compound (e.g., a fatty acid or fatty acid analog) which binds to FATP6 with high specificity.

Targeting vehicles specific to the heart-specific protein FATP6 have *in vivo* (e.g., therapeutic and diagnostic) applications. For example, an antibody which specifically binds to FATP6 can be conjugated to a drug to be targeted to the heart (e.g., a cardiac glycoside to treat congestive heart failure, or β-adrenergic agents, sodium channel blockers or calcium channel blockers to treat arrhythmias). A substance (e.g., a radioactive substance) which can be detected (e.g., a label) *in vivo* can also be linked to a targeting vehicle which specifically binds to a heart-specific protein such as FATP6, and the conjugate can be used as a labeling agent to identify cardiac muscle cells.

Targeting vehicles specific to FATP6 find further applications *in vitro*. For example, an FATP6-specific targeting vehicle, such as an antibody (a polyclonal preparation or monoclonal) which specifically binds to FATP6, can be linked to a substance which can be used as a stain for a tissue sample (e.g., horseradish peroxidase) to provide a method for the identification of cardiac muscle in a sample, as can be used in embryology studies, for example.

In a similar manner, an agent can be directed to the liver of a mammal, as FATP5 is expressed in liver but not in other tissue types. A targeting vehicle which specifically binds to FATP5 can be conjugated to a drug for delivery of the drug to the liver, such as a drug to treat hepatitis, Wilson's disease, lipid storage diseases and liver cancer. As with targeting vehicles specific to FATP6, targeting vehicles specific to FATP5 can be used in studying tissue samples *in vitro*.

The invention also relates to compositions comprising a modulator of FATP function. The term "modulate" as used herein refers to the ability of a molecule to alter the function of another molecule. Thus, modulate could mean, for example, inhibit, antagonize, agonize, upregulate, downregulate, induce, or suppress. A modulator has the capability of altering function of its target. Such alteration can be accomplished at any stage of the transcription, translation, expression or function of the protein, so that, for example, modulation of a target gene can be accomplished by modulation of the DNA or RNA encoding the protein, and the protein itself.

Antagonists or agonists (inhibitors or enhancers) of the FATPs of the invention, antibodies that bind a FATP, or mimetics of a FATP can be employed in combination with a non-sterile or sterile carrier or carriers for use with cells, tissues or organisms, such as a pharmaceutical carrier suitable for administration to a mammalian subject. Such compositions comprise, for instance, a media additive or a therapeutically effective amount of an inhibitor or enhancer compound to be identified by an assay of the invention and a pharmaceutically acceptable carrier or excipient. Such carriers may include, but are not limited to, saline, buffered saline, dextrose, water, ethanol, surfactants, such as glycerol, excipients such as lactose and combinations thereof. The formulation can be chosen by one of ordinary skill in the art to suit the mode of administration. The chosen route of administration will be influenced by the predominant tissue or organ location of the FATP whose function is to be inhibited or enhanced. For example, for affecting the function of FATP4, a preferred administration can be oral or through a tube inserted into the stomach (e.g., direct stomach tube or nasopharyngeal tube), or through other means to accomplish delivery to the small

intestine. The invention further relates to diagnostic and pharmaceutical packs and kits comprising one or more containers filled with one or more of the ingredients of the aforementioned compositions of the invention.

Compounds of the invention which are FATPs, FATP fusion proteins, FATP mimetics, FATP gene-specific antisense poly- or oligonucleotides, inhibitors or enhancers of a FATP may be employed alone or in conjunction with other compounds, such as therapeutic compounds. The pharmaceutical compositions may be administered in any effective, convenient manner, including administration by topical, oral, anal, vaginal, intravenous, intraperitoneal, intramuscular, subcutaneous, intranasal, transdermal or intradermal routes, among others. In therapy or as a prophylactic, the active agent may be administered to an individual as an injectable composition, for example as a sterile aqueous dispersion, preferably isotonic.

Alternatively, the composition may be formulated for topical application, for example, in the form of ointments, creams, lotions, eye ointments, eye drops, ear drops, mouthwash, impregnated dressings and sutures and aerosols, and may contain appropriate conventional additives, including, for example, preservatives, solvents to assist drug penetration, and emollients in ointments and creams. Such topical formulations may also contain compatible conventional carriers, for example cream or ointment bases, and ethanol or oleyl alcohol for lotions.

In addition, the amount of the compound will vary depending on the size, age, body weight, general health, sex, and diet of the host, and the time of administration, the biological half-life of the compound, and the particular characteristics and symptoms of the disorder to be treated. Adjustment and manipulation of established dose ranges are well within the ability of those of skill in the art.

A further aspect of the invention is a method to identify a polymorphism, or the presence of an alternative or variant allele of a gene in the genome of an organism (of interest here, genes encoding FATPs). As used herein, polymorphism refers to the occurrence of two or more genetically determined alternative sequences or alleles in a population. A polymorphic locus may be as small as a base pair. Polymorphic markers

include restriction fragment length polymorphisms, variable number of tandem repeats (VNTR's), hypervariable regions, minisatellites, dinucleotide repeats, trinucleotide repeats, tetranucleotide repeats, simple sequence repeats, and insertion elements such as Alu. The first identified alleleic form, or the most frequently occurring form can be arbitrarily designated as the reference (usually, "wildtype") form, and other allelic forms are designated as alternative (sometimes, "mutant" or "variant"). Dipolid organisms may be homozygous or heterozygous for allelic forms.

An "allele" or "allelic sequence" is an alternative form of a gene which may result from at least one mutation in the nucleotide sequence. Alleles may result in altered mRNAs or polypeptides whose structure or function may or may not be altered. Any given gene may have none, one, or many allelic forms (polymorphism). Common mutational changes which give rise to alleles are generally ascribed to natural deletions, additions, or substitutions of nucleotides. Each of these types of changes may occur alone, or in combination with the others, one or more times in a given sequence.

Several different types of polymorphisms have been reported. A restriction fragment length polymorphism (RFLP) is a variation in DNA sequence that alters the length of a restriction fragment (Botstein *et al., Am. J. Hum. Genet. 32*:314-331 (1980)). The restriction fragment length polymorphism may create or delete a restriction site, thus changing the length of the restriction fragment. RFLPs have been widely used in human and animal genetic analyses (see WO 90/13668; WO 90/11369; Donis-Keller, *Cell 51*:319-337 (1987); Lander et al., *Genetics 121*:85-99 (1989)). When a heritable trait can be linked to a particular RFLP, the presence of the RFLP in an individual can be used to predict the likelihood that the individual will also exhibit the trait.

Other polymorphisms take the form of short tandem repeats (STRs) that include tandem di-, tri- and tetra-nucleotide repeated motifs. These tandem repeats are also referred to as variable number tandem repeat (VNTR) polymorphisms. VNTRs have been used in identity and paternity analysis (US 5,075,217; Armour *et al., FEBS Lett. 307*:113-115 (1992); Horn *et al.*, WO 91/14003; Jeffreys, EP 370,719), and in a large number of genetic mapping studies.

Other polymorphisms take the form of single nucleotide variations between individuals of the same species. Such polymorphisms are far more frequent than RFLPs, STRs (short tandem repeats) and VNTRs (variable number tandem repeats). Some single nucleotide polymorphisms occur in protein-coding sequences, in which case, one of the polymorphic forms may give rise to the expression of a defective or other variant protein and, potentially, a genetic disease. Other single nucleotide polymorphisms occur in noncoding regions. Some of these polymorphisms may also result in defective protein expression (e.g., as a result of defective splicing). Other single nucleotide polymorphisms have no phenotypic effects.

Many of the methods described below require amplification of DNA from target samples and purification of the amplified products. This can be accomplished by PCR, for instance. See generally, *PCR Technology, Principles and Applications for DNA Amplification* (ed. H.A. Erlich), Freeman Press, New York, NY, 1992; *PCR Protocols: A Guide to Methods and Applications* (eds. Innis, et al.), Academic Press, San Diego, CA, 1990; Mattila *et al., Nucleic Acids Res. 19*:4967 (1991); Eckert *et al., PCR Methods and Applications 1*:17 (1991); *PCR* (eds. McPherson *et al.*, IRS Press, Oxford); and US 4,683,202.

Other suitable amplification methods include the ligase chain reaction (LCR) (see Wu and Wallace, *Genomics 4*:560 (1989); Landegren *et al., Science 241*:1077 (1988)), transcription amplification (Kwoh *et al., Proc. Natl. Acad. Sci. USA 86*:1173 (1989), self-sustained sequence replication (Guatelli *et al., Proc. Natl. Acad. Sci. USA 87*:1874 (1990), and nucleic acid based sequence amplification (NASBA). The latter two amplification methods involve isothermal reactions based on isothermal transcription, which produce both single stranded RNA (ssRNA) and double stranded DNA (dsDNA) as the amplification products in a ratio of about 30 or 100 to 1, respectively.

Another aspect of the invention is a method for detecting a variant allele of a human FATP gene, comprising preparing amplified, purified FATP DNA from a reference human and amplified, purified, FATP DNA from a "test" human to be

compared to the reference as having a variant allele, using the same or comparable amplification procedures, and determining whether the reference DNA and test DNA differ in DNA sequence in the FATP gene, whether in a coding or a noncoding region, wherein, if the test DNA differs in sequence from the reference DNA, the test DNA comprises a variant allele of a human FATP gene. The following is a discussion of some of the methods by which it can be determined whether the reference FATP DNA and test FATP DNA differ in sequence.

Direct Sequencing. The direct analysis of the sequence of variant alleles of the present invention can be accomplished using either the dideoxy chain termination method or the Maxam and Gilbert method (see Sambrook *et al., Molecular Cloning: A Laboratory Manual,* 2nd ed., Cold Spring Harbor Press, New York 1989; Zyskind *et al., Recombinant DNA Laboratory Manual*, Acad. Press, 1988)).

Denaturing Gradient Gel Electrophoresis. Amplification products generated using the polymerase chain reaction can be analyzed by the use of denaturing gradient gel eletrophoresis. Different alleles can be identified based on the different sequence-dependent strand dissociation properties and electrophoretic migration of DNA in solution (chapter 7 in Erlich, ed. *PCR Technology, Principles and Applications for DNA Amplification*, W.H. Freeman and Co., New York, 1992).

Single-strand Conformation Polymorphism Analysis. Alleles of target sequences can be differentiated using single-strand conformation polymorphism analysis, which identifies base differences by alteration in electrophoretic migration of single stranded PCR products, as described in Orita *et al., Proc. Natl. Acad. Sci. USA 86*:2766-2770 (1989). Amplified PCR products can be generated as described above, and heated or otherwise denatured, to form single-stranded amplification products. Single-stranded nucleic acids may refold or form secondary structures which are partially dependent on the base sequence. The different electrophoretic mobilities of single-stranded amplification products can be related to base-sequence differences between alleles of target sequences.

Detection of Binding by Protein That Binds to Mismatches. Amplified DNA comprising the FATP gene or portion of the gene of interest from genomic DNA, for example, of a normal individual is prepared, using primers designed on the basis of the DNA sequences provided herein. Amplified DNA is also prepared, in a similar manner, from genomic DNA of an individual to be tested for bearing a distinguishable allele. The primers used in PCR carry different labels, for example, primer 1 with biotin, and primer 2 with ^{32}P. Unused primers are separated form the PCR products, and the products are quantitated. The heteroduplexes are used in a mismatch detection assay using immobilized mismatch binding protein (MutS) bound to nitrocellulose. The presence of biotin-labeled DNA wherein mismatched regions are bound to the nitrocellulose via MutS protein, is detected by visualizing the binding of streptavidin to biotin. See WO 95/12689. MutS protein has also been used in the detection of point mutations in a gel-mobility-shift assay (Lishanski, A. *et al., Proc. Natl. Acad. Sci. USA 91*:2674-2678 (1994)).

Other methods, such as those described below, can be used to distinguish a FATP allele from a reference allele, once a particular allele has been characterized as to DNA sequence.

Allele-specific probes. The design and use of allele-specific probes for analyzing polymorphims is described by e.g., Saiki *et al., Nature 324*:163-166 (1986); Dattagupta, EP 235,726, Saiki, WO 89/11548. Allele-specific probes can be designed so that they hybridize to a segment of a target DNA from one individual but do not hybridize to the corresponding segment from another individual due to the presence of different polymorphic forms in the respective segments from the two individuals. Hybridization conditions should be sufficiently stringent that there is a significant difference in hybridization intensity between alleles, and preferably an essentially binary response, whereby a probe hybridizes to only one of the alleles. Some probes are designed to hybridize to a segment of target DNA such that the polymorphic site aligns with a central position (e.g., in a 15-mer at the 7 position; in a 16-mer, at either the 8 or

9 position) of the probe. This design of probe achieves good discrimination in hybridization between different allelic forms.

Allele-specific probes are often used in pairs, one member of a pair showing a perfect match to a reference form of a target sequence and the other member showing a perfect match to a variant form. Several pairs of probes can then be immobilized on the same support for simultaneous analysis of multiple polymorphisms within the same target sequence.

Allele-specific Primers. An allele-specific primer hybridizes to a site on target DNA overlapping a polymorphism, and only primes amplification of an allelic form to which the primer exhibits perfect complementarity. See Gibbs, *Nucleic Acid Res.* *17*:2427-2448 (1989). This primer is used in conjunction with a second primer which hybridizes at a distal site. Amplification proceeds from the two primers, resulting in a detectable product which indicates the particular allelic form is present. A control is usually performed with a second pair of primers, one of which shows a single base mismatch at the polymorphic site and the other of which exhibits perfect complementarity to a distal site. The single-base mismatch prevents amplification and no detectable product is formed. The method works best when the mismatch is included in the 3'-most position of the oligonucleotide aligned with the polymorphism because this position is most destabilizing to elongation from the primer (see, e.g., WO 93/22456).

Gene Chips. Allelic variants can also be identified by hybridization to nucleic acids immobilized on solid supports (gene chips), as described, for example, in WO 95/11995 and U.S. Patent No. 5,143,854, both of which are incorporated herein by reference. WO 95/11995 describes subarrays that are optimized for detection of a characterized variant allele. Such a subarray contains probes designed to be complementary to a second reference sequence, which is an allelic variant of the first reference sequence.

The present method is illustrated by the following examples, which are not intended to be limiting in any way.

EXAMPLES

Materials and Methods

The following Materials and Methods were used in the work described in Examples 1-5.

Sequence Alignment of FATP Clones. The DNA sequence for mouse FATP1 was obtained from the National Center for Biotechnology Information nonredundant database. cDNAs for mmFATP2, 3, 4, and 5 were obtained by screening mouse expression libraries (purchased from GIBCO/BRL) with probes derived from the cloned expressed sequence tags (ESTs) (Research Genetics, Huntsville, AL). Full-length clones were obtained for mmFATP2 and 5 and partial sequences for mmFATP3 and 4. The sequences described herein have been deposited in the GenBank database (Accession Nos. FATP2, AF072760; FATP3, AF072759; FATP4, AF072758; FATP5, AF072757).

Neither FATP2 nor FATP5 contains an in-frame stop codon upstream of the putative initiator methionine; initiator methionines were assigned by homology with that in mmFATP1 and by the presence of a signal sequence immediately after it. The *Mycobacterium tuberculosis, Caenorhabditis elegans*, and *Saccharomyces cerevisiae* sequences were present in the dbEST database as part of the sequencing projects for these organisms. Sequences were aligned utilizing a ClustalX algorithm and the resulting alignment exported to SeqVu. Homologous amino acid substitutions are boxed in Figure 1 and were determined using the Dayhoff 250 method with a 50% homology cutoff.

Cell Transfection and LCFA Uptake. COS cells were cotransfected using the DEAE-dextran method with the mammalian expression vector pCDNA 3.1 (Invitrogen) expressing the gene for CD2 (pCDNA-CD2) in combination with either a pCDNA 3.1 or pCMVSPORT2 (GIBCO/BRL) expression vector containing one of the murine or

nematode *FATP* genes (*pCDNA-mmFATP1*, *pCDNA-FATP2*, *pCMVSPORT-FATP5*, *pCDNA-ceFATPb*). Two days after transfection, cells were assayed for CD2 expression with a phycoerythrin-coupled anti-CD2(PE-CD2) monoclonal antibody (PharMingen), and fatty acid uptake was assayed with a BODIPY-labeled fatty acid analogue (Molecular Probes). Briefly, cells were washed twice with PBS (phosphate buffered saline) and stained with PE-CD2 at 4°C for 30 min in PBS containing 10% fetal calf serum. They were then washed three times with PBS/fetal calf serum for 5 min followed by an incubation for 2 min at 37°C in fatty acid uptake solution, which contained 0.1 μM BODIPY-FA and 0.1% fatty acid-free BSA (bovine serum albumin) in PBS (Schaffer, J.E. & Lodish, H.F. (1994) *Cell 79:*427-436). After 2 min, the cells were washed four times with ice-cold PBS/0.1% BSA. The cells were then removed from the plates with PBS containing 5 mM EDTA and resuspended in PBS containing 10% fetal calf serum and 10 mM EDTA. PE-CD2 and BODIPY-FA fluorescence were measured using a FACScan (Becton Dickinson). COS cells were gated on forward scatter (FSC) and side scatter (SS). Cells exhibiting more than 300 CD2 fluorescence units (dsim) representing 15% of all cells were deemed CD2 positive and their BODIPY-FA fluorescence was quantitated.

E. coli-Based LCFA Uptake Assay. The full-length coding region of mtFATP and a control protein, the mammalian transcription factor TFE3, were subcloned into the inducible, prokaryotic expression vector pET (Novagen). Expression was induced with 1 mM isopropyl β-D-thiogalactoside (IPTG) for 1 hour, or cells were left uninduced. Cells were washed in PBS/0.1% BSA and resuspended in 1 ml PBS/0.1% BSA containing 0.1 μM [^{3}H]palmitate (NEN) at 37°C. Uptake was stopped after the indicated incubation time by transferring the cells onto filter paper using a cell harvester (Brandel, Bethesda, MD). Filters were washed extensively with ice-cold PBS/0.1% BSA, and [^{3}H]palmitate was quantitated by scintillation counting.

Northern Blots. Northern blot analysis of murine FATP expression was done using poly(A) mRNA blots (Clontech). Probes of each of the FATPs were derived from the 3' untranslated regions of each gene and were <60% identical in sequence. Probes

were labeled by random priming (Boehringer Mannheim) and hybridized at 65°C. Blots were extensively washed in 0.2% SSC/0.1% SDS at 65°C.

Generation of Phylogenetic Trees. Complete and partial sequences for *FATP* genes from human, rat, mouse, puffer fish, *Drosophila melanogaster*, *C. elegans*, *S. cerevisiae*, and *M. tuberculosis* were aligned using ClustalX. A homologous region of 48 amino acids (residues 472-519 in mmFATP1) from all of the genes was used to determine phylogenetic relationship within ClustalX. Based on these data a phylogenetic tree was generated using Tree View PPC (Figure 5).

Nomenclature. It is proposed that the *FATP* genes be given a species specific prefix (mm, *Mus musculus*; hs, *Homo sapiens*; mt, *M. tuberculosis*; dm, *D. melanogaster*; ce, *C. elegans,* sc, *S. cerevisiae*) and numbered such that mammalian homologues in different species share the same number but differ in their prefix. Since the two *C. elegans* genes cannot be paired with a specific human or mouse FATP, they have been designated *ceFATPa* and *ceFATPb*.

Example 1: Identification of Novel Mammalian FATPs

The National Center for Biotechnology Information EST database was screened, using the mouse FATP protein sequence (mmFATP1), to identify novel FATPs. This strategy led to the identification of more than 50 murine EST sequences which could be assembled into five distinct contiguous DNA sequences (contigs). One contig was identical to the previously cloned FATP, which has been renamed FATP1. Another, which has been renamed FATP2, is the murine homologue of a rat gene previously identified by others as a very long chain acyl-CoA synthase (Uchiyama, A., Aoyama, T., Kamijo, K., Uchida, Y., Kondo, N., Orii, T. & Hashimoto, T. (1996) *J. Biol. Chem. 271*:30360-30365). The other three contigs represented novel genes (*FATP3*, *4*, and *5*). Full-length clones for *FATP2* and *FATP5* and nearly complete sequences for *FATP3* and *4* (Figure 1) were obtained by screening cDNA libraries made from mouse day 10.5 embryos and adult liver. Also identified were human homologues for each of the murine genes in the EST database. A sixth human gene was also identified; whether

this gene is also present in the mouse will require additional studies. Map positions are given in Tables 2 and 3.

The genetic loci for all of the human genes, with the exception of FATP5 which was already mapped as an unknown EST, were determined using the radiation hybrid panels. The map positions given below show the distance (in centiRays) from the closest framework marker. As a guideline, there are approximately 300kb/cR.

Table 2. Mapping Data for Human Genes

hsFATP1	Chromosome Chr19
	places 13.35 cR from WI-6344 (lod>3.0)
hsFATP2	Chromosome Chr15
	places 4.92 cR from D15S126 (lod>3.0)
hsFATP3	Chromosome Chr1
	places 13.24 cR from WI-2862 (lod>3.0)
hsFATP4	Chromosome Chr9
	places 7.80 cR from WI-9685 (lod>3.0)
hsFATP5	unknown EST previously mapped to near D19S418
hsFATP6	Chromosome Chr5
	places 1.41 cR from WI-4907 (lod>3.0)

The mouse map is an internal backcross panel consisting of 188 mouse backcross DNA's plus 4 controls (B6, Spretus, F1, Water). The backcross was constructed by crossing B6 by Spretus animals and then crossing those F1's back to B6. Mapping is accomplished by taking advantage of recombinational events during meiosis, and the use of PCR primers to detect the differences (by size or re-annealing events) at any given locus between the B6 and Spretus allele.

For the purposes of mapping, a novel set of primers (gene of interest) is used to amplify from all 188 DNA's and then typed as being a B6 ("B") or a Spretus ("S"). This

string of B's and S's is entered into the Map Manager program, which does a best fit calculation by comparing the string of 188 typings from the gene of interest to all loci already extant in the panel, for all 20 chromosomes. The gene of interest is then assigned to a particular area on a particular chromosome according to a number of parameters, including the minimalization of double cross-overs, and the highest LOD scores. Indicated in Table 3 are distances to the closest markers on either side of the FATP locus.

Table 3. Mapping Data for Mouse Genes

mmFATP1	Chromosome 8 places 2.82 cM from D8Mit132 (lod 43.4) and 1.81 cM from D8Mit74 (lod 43.5)
mmFATP2	Chromosome 2 places 1.29 cM from D2Mit258 (lod 47.9) and 1.75 cM from D2NDS3 (lod 44.9)
mmFATP3	Chromosome 3 places 2.54 cM from D3Mit22 (lod 29.5) and 19.62 cM from D3Mit42 (lod 13.6)
mmFATP4	Chromosome 2 places 13.78 cM from D2Mit1 (lod 22.9) and 3.85 cM from D2Mit65 (lod 41.9)
mmFATP5	Chromosome 7 places 7.28 cM proximal of D7Mit21 (lod 28.3)

Example 2: Assessment of Function

The ability of the newly identified mouse genes to function as fatty acid transporters was assessed using a fluorescence-activated cell sorting-based assay. COS cells were transiently cotransfected with expression vectors encoding the cell surface

protein CD2 and either mmFATP1, mmFATP2, or mmFATP5, respectively. Two days after transfection, COS cells were stained with an antibody to CD2 and then incubated with a BODIPY-labeled fatty acid [BODIPY-FA, (Schaffer, J.E. & Lodish, H.F. (1994) *Cell 79:*427-436)]. The cells were then washed extensively, lifted off the dish, and analyzed by fluorescence-activated cell sorting. As judged by the number of CD2-positive cells, the transfection efficiency was approximately 20-30%. Fatty acid uptake was quantitated in the transiently transfected COS cells by measuring the BODIPY-FA fluorescence of the CD2-positive cells. Expression of CD2 had no effect on fatty acid uptake as shown by the finding that COS cells expressing only the transfected CD2 cDNA (CD2-positive) had the same low level of BODIPY-FA uptake as did untransfected (CD2-negative) control cells (Figure 2A, control). In COS cells cotransfected with CD2 and mmFATP1, mmFATP2, or mmFATP5, uptake of BODIPY-FA by the transfected (CD2-positive) cells was increased between 15- to 90-fold over control (CD2 cDNA only) cells (Figures 2A-2D).

Example 3: Expression Patterns of Murine FATPs

Expression patterns of members of the murine *FATP* gene family were characterized by Northern blot analysis; to avoid cross-hybridization, the probes used were from the 3' untranslated region of these genes, which are less than 60% identical in sequence. The expression pattern of FATP1 agrees with that previously found (Schaffer, J.E. & Lodish, H.F. (1994) *Cell 79:*427-436). Here, expression was seen primarily in heart and kidney. FATP2 is expressed almost exclusively in liver and kidney, which corresponds to the reported tissue distribution of the rat homologue [very long chain acyl-CoA (VLACS)] as assessed by Western blotting (Uchiyama, A., Aoyama, T., Kamijo, K., Uchida, Y., Kondo, N., Orii, T. & Hashimoto, T. (1996) *J. Biol. Chem. 271:*30360-30365). FATP3 is present in lung, liver, and testis. FATP5 is expressed only in liver and cannot be detected in other tissues even when the blot is overexposed. The human homologue of FATP5 is also liver specific and is not expressed in a wide array of other tissues tested, including fetal liver.

Example 4: FATPs Are Evolutionarily Conserved

The EST database was searched, using sequences conserved among the five murine FATP genes, for *FATP* genes in other organisms. Two homologues were found in *C. elegans* and one in *M. tuberculosis*. One of the *C. elegans* genes was cloned from a cDNA library and expressed in COS cells, as described for the murine FATPs. Overexpression of the nematode FATP resulted in a 15-fold increase of BODIPY-FA uptake compared with control cells (Figure 3). The mycobacterial *FATP* gene was isolated from a phage library and assessed for its ability to facilitate fatty acid uptake. *E. coli* transformed with a prokaryotic, isopropyl β-D-thiogalactoside-inducible expression vector containing the mycobacterial *FATP* gene demonstrated a significant increase in the rate of [^{3}H]palmitate uptake after induction, compared with uninduced bacteria or *E. coli* transformed with a control protein (Figure 4). Novel *FATP* genes were also identified in *F. rubripes* (puffer fish) and *D. melanogaster*.

Example 5: Phylogenetic Tree of FATPs

Faergeman *et al.* (Faergeman, N.J., DiRusso. C.C., Elberger, A., Knudsen, J. & Black, P. N. (1997) *J. Biol. Chem. 272*:8531-8538) identified three regions of very strong conservation between the *scFATP* and *mmFATP1* genes. The sequences of the FATPS were compared over a 311-amino acid FATP "signature sequence" which includes these conserved regions corresponding to amino acids 246-557 in mmFATP1 (underlined in Figure 1). When compared with the National Center for Biotechnology Information nonredundant database, only one region of the "FATP signature sequence" shows significant homology to other proteins. This small stretch of amino acids (underlined in Fig. 1) is an AMP-binding motif found in a multitude of other proteins, such as acyl-CoA synthase, several CoA lipases, and gramicidin S synthetase component II (Schaffer, J.E. & Lodish, H.F. (1994) *Cell 79:*427-436). The relevance of this motif to fatty acid transport is unclear. Other highly conserved regions among the FATPs, including long stretches of amino acids >90% identical from mycobacteria to humans, are not found in any other class of proteins. A 48-amino acid segment of the

FATP signature sequence was used to construct a phylogenetic tree (Figure 5). Each of the human and mouse genes form their own branch; hsFATP6, which as yet has no murine homologue, is most closely related to hsFATP3 and mmFATP3. As expected, rnVLACS is closer in sequence to mmFATP2 than to hsFATP2. The *FATP* genes of invertebrates i.e., *C. elegans* and *D. melanogaster*, are most closely related to each other. Surprisingly, the mycobacteral gene is more closely related to the human and mouse *FATP5* genes than to the FATPs of any of the lower organisms. Whether this reflects coevolution of the mycobacterial and human genes awaits further study.

Materials and Methods

The following materials and methods were used in the work described in Examples 6-10.

Isolation of full-length human FATP1 and 4

Full-length clones encoding human FATP1 and human FATP4 were identified by searching databases for sequences similar to murine FATP1-5 coding regions using the BlastX algorithm (Altschul *et al., J. Mol. Biol. 215*: 403-410, 1990).

A concatamer of nucleotide sequences comprising the coding sequences of mmFATP1 (Genbank Accession U15976), mmFATP2, mmFATP3 (SEQ ID NO:6), mmFATP4 (SEQ ID NO:8) and mmFATP5 (SEQ ID NO:10) was used to search the Millennium database using the BLASTX algorithm. Sequences with a score >150 were evaluated for whether they represented known FATP coding sequences.

Human clones with similarity to the 5' end of murine FATP sequences were sequenced completely. Clones encoding full-length human FATP1 were obtained from a heart cDNA library constructed in the mammalian expression vector pMET7 (Tartaglia *et al., Cell, 83*: 1263-1271, 1995). Clones encoding full-length human FATP4 were obtained from a spleen cDNA library constructed in the mammalian expression vector pMET7.

Isolation of full-length human FATP6

Several clones encoding human FATP6 were identified by searching public databases as described above. Five clones were analyzed further by restriction digestion and DNA sequencing. One of these clones (Genbank Accession # AA412064) appeared to be full-length and its entire insert was sequenced.

DNA Sequence Analysis

Sequences were aligned with the DNAStar program using the Clustal method. Hydrophobicity plots were generated with DNA Strider using the Kyte Doolittle method.

In situ hybridization

Tissues were collected from 8 week old C57/B16 mice. Tissues were fresh frozen, cut on a cryostat at 10 μm thickness and mounted on Superfrost Plus slides (VWR). Sections were air dried for 20 minutes and then incubated with ice cold 4% paraformaldehyde (PFA)/phosphate buffered saline (PBS) for 10 minutes. Slides were washed 2 times 5 minutes with PBS, incubated with 0.25% acetic anhydride/1 M triethanolamine for 10 minutes, washed with PBS for 5 minutes and dehydrated with 70%, 80%, 95% and 100% ethanol for 1 minute each. Sections were incubated with chloroform for 5 minutes. Hybridizations were performed with ^{35}S-radiolabeled ($5x10^7$ cpm/ml) cRNA probes generated from the 3' untranslated regions of mouse FATPs by PCR followed by *in vitro* transcription in the presence of 50% formamide, 10% dextran sulfate, 1x Denhardt's solution, 600 mM NaCl, 10 mM DTT, 0.25% SDS and 10 μg/ml tRNA for 18 hours at 55°C. After hybridization, slides were washed with 10 mM Tris-HCl pH 7.6, 500 mM NaCl, 1 mM EDTA (TNE) for 10 minutes, incubated in 40 μg/ml RNase A in TNE at 37°C for 30 minutes, washed in TNE for 10 minutes, incubated once in 2x SSC at 60°C for 1 hour, once in 0.2x SSC at 60°C for 1 hour, once in 0.2x SSC at 65°C for 1 hour and dehydrated with 50%, 70%, 80%, 90% and 100% ethanol. Localization of mRNA transcripts was detected by dipping slides in Kodak

NBT-2 photoemulsion and exposing for 7 days at 4°C, followed by development with Kodak Dektol developer. Slides were counter stained with haematoxylon and eosin and photographed. Controls for the in situ hybridization experiments include the use of a sense probe which showed no signal above background in all cases.

Northern Blotting

Human mRNA blots were obtained from Invitrogen or Clontech. PCR fragments from the 3' untranslated regions of human FATPs were used as probes. Blots were probed with ^{32}P-labeled DNA probes using the Rapid-Hyb buffer (Amersham) according to the manufacturer's instructions.

Cell transfection and LCFA uptake. COS cells were cotransfected, using lipofectamine (GIBCO BRL) according to the manufacturer's instructions, with the mammalian expression vector pCDNA3.1 (Invitrogen) expressing the gene for CD2 in combination with a pMET7 expression vector (Tartaglia *et al., Cell*, 83:1263-1271, 1995) containing hsFATP1 (pMET7-hsFATP1) or hsFATP4 (pMET7-hsFATP4) or pMET7 alone. Two days after transfection, cells were assayed for CD2 expression with a phycoerythrin-coupled anti-CD2 (PE-CD2) monoclonal antibody (PharMingen), and fatty acid uptake was assayed with a BODIPY-labeled fatty acid analog (Molecular Probes) as described above.

Example 6: Determination of Expression of mmFATPs

mmFATP4, and to lesser extent mmFATP2, are expressed at high levels in the brush border layer of the small intestine.

Cell transfection and LCFA uptake. COS cells were cotransfected, using lipofectamine (GIBCO BRL) according to the manufacturer's instructions, with the mammalian expression vector pCDNA3.1 (Invitrogen) expressing the gene for CD2 in combination with a pMET7 expression vector (Tartaglia *et al., Cell*, 83:1263-1271, 1995) containing hsFATP1 (pMET7-hsFATP1) or hsFATP4 (pMET7-hsFATP4) or pMET7 alone. Two days after transfection, cells were assayed for CD2 expression with

a phycoerythrin-coupled anti-CD2 (PE-CD2) monoclonal antibody (PharMingen), and fatty acid uptake was assayed with a BODIPY-labeled fatty acid analog (Molecular Probes) as described above.

Absorption of dietary fat requires transport of free fatty acids across the apical membrane of epithelial cells in the small intestine. Previous studies suggested that this transport is protein-mediated; however, the transport protein had not yet been identified. In situ hybridization was performed on each of the three regions of the small intestine -- duodenum, jejunum and ileum -- as well as the colon, using probes from the 3' untranslated regions of mmFATP1, mmFATP2, mmFATP3, mmFATP4 and mmFATP5, to determine whether any of the mouse FATPs are expressed in the small intestine. It was expected that a protein involved in fatty acid absorption would be expressed in the epithelial cells of the small intestine, but absent from the colon.

Expression of mmFATPs in the jejunum was identical to that in the ileum in all cases. High levels of mmFATP4 mRNA were present in the epithelial cells of the jejunum and ileum, and lower, but significant, amounts were detected in the epithelial cells of the duodenum. Significantly, FATP4 mRNA was absent from other cell types of the small intestine and no FATP4 mRNA could be detected in any of the cells of the colon. FATP2 mRNA was present in the epithelial cells of the duodenum at a level similar to that of FATP4, but was present at lower levels in the jejunum and ileum. No signals above background were detected for mmFATP1, mmFATP3 and mmFATP5 in any of the intestinal tissues. mmFATP3 and FATP5 were clearly detectable by in situ hybridization in adult liver and mmFATP1 could be detected in a variety of tissues on a whole embryo in situ, indicating that the FATP1, 3, and 5 probes were working.

mmFATP4 expression is predominant in the small intestine compared to the other organs of the mouse embryo. In the small intestine, FATP4 expression is limited to differentiated enterocytes, while no signal is detected in the connective tissue or the undifferentiated epithelial cells in the crypts. Differentiated enterocytes are known to be the cells that mediate the uptake of fatty acids. FATP4 is specifically and strongly expressed in the epithelial cells of adult murine duodenum and ileum but not colon.

Other FATPs, such as FATP5, are not expressed in the small intestine. Thus, FATP4 is the major FATP in the mouse small intestine. Given its high level of expression, it is likely that FATP4, and to a lesser extent FATP2, play an important role in the absorption of fatty acids.

mmFATP2, and mmFATP5 are expressed in hepatocytes

Northern analysis of mmFATP2, mmFATP3, mmFATP4 and mmFATP5 showed expression in the liver. To determine whether these proteins are present in hepatocytes or other cells types present in liver homogenates, in situ hybridizations were performed. mmFATP2, and mmFATP5 mRNA was clearly present in hepatocytes, and was not concentrated in other cell types such as endothelial cells or macrophages. No signal above background was detected for mmFATP1 in any of the cell types in the liver, consistent with the results of the Northern blotting.

Example 7: Isolation and Sequence Analysis of Full-length Human FATP1 and Full-length Human FATP4

To identify human cDNA clones encoding FATP family members, Millennium databases were searched for sequences similar to murine FATP1-5 coding regions. Two clones were analyzed in detail; inspection of the entire DNA sequence of these two clones showed that they encode the human orthologs of mmFATP1 and mm FATP4, respectively. These two clones were designated hsFATP1 and hsFATP4, and their DNA and predicted protein sequences are shown in Figures 44A-44C and 45, and 50A-50C and 51. hsFATP1 is predicted to encode a 646 amino acid, 71 kD protein with multiple membrane-spanning domains (Figure 28A). HsFATP4 is predicted to encode a 643 amino acid, 72 kD protein with multiple membrane spanning domains (See Figure 29A). A comparison of the DNA sequences of mouse and human FATP1 and mouse and human FATP4 (Figures 30A-30B and 31A-31B) shows that the mouse and human orthologs are 85% (FATP1) and 87% (FATP4) identical to each other within the coding sequences given in these figures. At the amino acid level, hsFATP1 and hsFATP4 are

~90% identical to their respective mouse orthologs within the coding region shown in these figures (Figures 32 and 33). The sequence identities between mouse and human FATP1 and FATP4 are considerably higher than the ones observed between different FATP family members within one species (~40%-60%) and are present in the N-terminal part of the protein, a region that is poorly conserved between different FATP family members. This high degree of sequence conservation clearly demonstrates that the newly identified human FATPs are orthologs of mouse FATP1 and FATP4 rather than novel FATP family members.

Table 4 is an identity/similarity matrix comparing the amino acid sequences of FATP1 and 4 from human and mouse. This shows that the gene whose sequence is shown in Figure 43A is indeed human FATP4, since it is 91% identical with the murine FATP4 but only 62% identical with the closest related human FATP, which is FATP1.

Table 4				
Identity/Similarity Matrix				
	hsFATP4	mmFATP4	hsFATP1	mmFATP1
hsFATP4	---	93.2	72.3	72.0
mmFATP4	91.0	---	71.2	71.1
hsFATP1	61.9	61.0	---	92.4
mmFATP1	60.7	59.6	89.5	---

Example 8: Isolation and Sequence Analysis of Full-length Human FATP6

A search of EST databases identified a set of overlapping human sequences that were similar to FATPs, but did not have a clear mouse ortholog. One of these EST clones was found to encode a full-length cDNA. The entire insert of this clone was sequenced and designated hsFATP6. The DNA and predicted protein sequences of hsFATP6 are shown in Figures 54A-54C and 55. HsFATP6 is predicted to encode a 619 amino acid, 70 kD protein with multiple membrane-spanning domains (Figure

35A). A comparison of the amino acid sequences of hsFATP6 with other human FATPs shows about 37% identity to either hsFATP1 or hsFATP4 (Figure 36). This degree of sequence identity is similar to what is observed between different mouse FATPs. The phylogenetic analysis described above clearly demonstrates that hsFATP6 is a member of the FATP family, but not an ortholog of any of the mouse FATPs. Comparisons were done with "ALIGN" (E. Myers and W. Miller, "Optimal Alignments in Linear Space," *CABIOS 4*:11-17 (1988) using standard settings.

Example 9: Tissue Distribution of Human FATPs

The tissue distribution of human FATPs was assessed by Northern blotting. Human FATP3 was expressed in a large variety of tissues. In contrast, human FATP5 was present at high levels in the liver, but was undetectable in all other tissues examined. Thus, both hsFATP3 and hsFATP5 recapitulate the expression pattern of their mouse orthologs (see above). HsFATP6 is a novel FATP with no mouse ortholog as yet. Northern blotting shows that hsFATP6 is expressed at high levels in the heart, but is undetectable in other tissues, including skeletal and smooth muscle. This tissue distribution suggests that human FATP6 performs an important role in energy metabolism in the heart; blocking FATP6-mediated fatty acid transport may therefore be beneficial for a number of heart diseases, e.g., ischemic heart disease.

To identify the major FATP expressed in the human small intestine, Northern blotting was performed on a blot containing mRNA from human stomach, jejunum, ileum, colon, rectum and lung. hsFATP5 and hsFATP6 were undetectable in any of these tissues. FATP5 is only expressed in liver and FATP6 only in heart. hsFATP2 was weakly expressed in the colon, and an even weaker signal was detectable in jejunum, ileum and lung lanes. hsFATP3 was expressed well in the lung, but was only weakly expressed in the other tissues tested. Importantly, no difference was seen in the expression of hsFATP3 between small intestine and stomach or colon, suggesting that the expression observed is not related to fatty acid absorption in the small intestine. hsFATP4 was clearly expressed in both jejunum and ileum; expression was

significantly lower in the colon and was absent in the stomach. This expression pattern is consistent with a major role for FATP4 in absorption of fatty acids in the human gut.

Example 10: Expression of hsFATP1 and hsFATP4 Promotes Transport of Fatty Acids

COS cells were cotransfected using lipofectamine with the mammalian expression vector pCDNA-CD2 in combination with one of the FATP-containing expression vectors (pMET7-hsFATP1 or pMET7-hsFATP4) or an insertless expression vector (pMET7, control) as described in Materials and Methods for Examples 6-10. COS cells were gated on forward scatter and side scatter. Cells exhibiting more than 400 CD2 fluorescence units representing ~30% of all cells were deemed CD2-positive. The percent of CD2-positive cells exhibiting a BODIPY-fluorescence of >300 is plotted for the three different vectors tested (Figure 37).

Example 11: Stable Expression of Human FATP4 in 293 Cells

Stable cell lines were generated as follows. A DNA fragment containing the entire hsFATP4 coding sequence as well as 100 nucleotides of 5′ and 50 nucleotides of 3' untranslated region was inserted into the vector pIRES-neo (Clontech) using standard cloning techniques. The resulting construct or a vector control (pIRES-neo) was transfected into 293 cells using the lipofectamine method (Gibco BRL) according to the manufacturer's directions. Cells that had taken up the DNA were selected with 1 mg/ml G418 (Gibco BRL). Single colonies were picked 1 to 2 weeks after transfection and grown in medium containing 0.8 mg/ml G418. Colonies were screened for the ability to take up fatty acids by measuring uptake of a fluorescently labeled fatty acid (BODIPY-FA). About 40 colonies transfected with the pIRES-neo containing FATP4 and ~20 colonies transfected with pIRES-neo control were analyzed. All 20 of the vector control clones showed amounts of BODIFY-FA uptake similar to each other and to untransfected 293 cells. In contrast, among the 40 FATP4 transfected clones, 3 had a 5- to 10-fold increased BODIPY-FA uptake compared to any of the vector controls, and a large number (~20) showed an approximately two-fold increase in BODIPY-FA levels.

This distribution is consistent with FATP4 conferring increased fatty acid uptake in these cells. One of the cell lines with the highest amount of BODIPY-FA uptake was selected to be used for measuring uptake of tritiated fatty acid.

The uptake of tritiated oleate over time by either FATP4 expressing or control cells was assayed over time. Expression of FATP4 increases the rate of fatty acid uptake by over 3-fold, demonstrating that FATP4 is, like the other FATPs, a functional fatty acid transporter (Figure 38).

Example 12: Immuno-staining with FATP4-Specific Antiserum

A polyclonal antiserum against the C-terminus of mmFATP4 was raised using a GST-fusion protein having mmFATP4-specific amino acid sequence 552-643 (AVASP...GEEKL). In western blot experiments, the purified antibody reacted strongly with a synthetic peptide matching the C-terminus of mmFATP4, but not with a corresponding region of mmFATP2, mmFATP3, or mmFATP5. The mmFATP4 specific polyclonal antiserum detects, in western blot experiments with enterocyte lysates from 3 different mice, a ~70 kDa protein, which is in accordance with mmFATP4's predicted molecular weight of 72 kDa. The binding is specific for mmFATP4, since it can be completely abolished by preincubation of the antiserum with the GST-fusion peptide used to raise the antibody.

Immunofluorescence experiments were performed using the anti-mmFATP4 antiserum on fresh frozen sections of murine small intestine. The antibody binding demonstrates strong expression of mmFATP4 in enterocytes, confirming the results of the in situ hybridization experiments. At higher magnifications it is apparent that mmFATP4 is expressed at the apical side of the enterocyte, indicating that the transporter is present in the brush border membrane, which is known to mediate the uptake of fatty acids from the intestinal lumen.

Immuno-electron microscopy studies were performed on fresh frozen murine intestinal cells. The gold particles used, appearing as black specks on the electron micrographs, indicate the subcellular localization of mmFATP4 to be on the microvilli

of the enterocyte. It can be seen from electron micrographs that mmFATP4 is localized exclusively in membranes, preferentially the apical plasma membrane, confirming that it is indeed a membrane protein.

Methods for Immunofluorescence and Immunogold Electron Microscopy

Unfixed mouse small intestine was washed with Hank's buffered salt solution containing 1 mM EDTA, infused with 2.3 M sucrose solution, and embedded in O.C.T., 4583 compound. The material was thick sectioned (15 μM - 40 μM). The sections were washed in PBS containing 1% BSA and 0.075% glycine to block non-specific binding. Primary and secondary antibodies were diluted in PBS with 10% FCS and incubated for 1h. The sections were mounted in 90% glycerol/PBS containing 1 mg/ml paraphenylinediamine, and examined with a Bio-Rad MRC 600 confocal, mounted on a Zeiss Axioscop.

For the immunogold labeling, the tissue was fixed with 2% paraformaldehyde in PBS for 10 minutes, after which it was cryoprotected by infiltration with 2.3 M sucrose in 0.1 M phosphate buffer (pH 7.4) containing 20% polyvinylpyrrolidone, and then mounted on aluminum cryo nails and frozen in liquid nitrogen (Tokuyasu, K.T., *J. Microscop. 143*:139-149, 1986). Ultrathin sections were collected on carbon/formvar-coated nickel grids. The primary antibody (anti-FATP4) was diluted in 10% FCS in PBS and incubated overnight at 4 C, followed by donkey anti-rabbit IgG-gold (12 nm) (Jackson Labs) for 1h. The sections were stained in 2% neutral uranyl acetate (20 minutes) and absorption stained with 2% uranyl acetate in 0.2% methylcellulose containing 3.2% polyvinyl alcohol. The sections were examined with a Philips EM 410 electron microscope.

Example 13: Inhibition of Fatty Acid Uptake Specific to FATP4 Demonstrated in Isolated Mouse Enterocytes

Phosphorothioate derivatives of the following oligonucleotides were synthesized:

FATP4-AS2	CCCCCACCAGAGAGGCTCC (SEQ ID NO:103)
FATP4-AS2MM	CCACCCCCGGAAAGCCTGC (SEQ ID NO:104)
FATP4-S2	GGAGCCTCTCTGGTGGGGG (SEQ ID NO:105)

FATP4 AS2 is the antisense oligo; it is designed to be complementary to the sequence extending from nucleotide 10 to nucleotide 28 of the mouse FATP4 coding sequence. FATP4-AS2MM is a control oligo; in the oligo every third nucleotide was changed creating mismatches; the overall nucleotide composition is identical to FATP4-AS2 (same number of G, A, T, C). FATP4-S2 is the sense control.

Enterocytes were isolated from the small intestine of mice and incubated for 48h in tissue culture (Figure 40) either without oligonucleotides (squares) or with 100 μM FATP4 specific sense (circles) or antisense (diamonds) oligonucleotides. The uptake over time of 25 μM oleate was then measured. While the FATP4 sense oligonucleotide did not significantly influence the uptake, the antisense oligonucleotide inhibited fatty acid uptake by ~50%.

The effect of either FATP4 sense, antisense or mismatch sequence oligonucleotides on the uptake of fatty acids was measured in enterocytes. Isolated enterocytes were incubated with increasing concentrations of FATP4 antisense oligonucleotides (solid bars in Figure 41), or a mismatch control oligonucleotide with identical nucleotide composition (stippled bars), or with 100 μM of the FATP4 sense-oligonucleotide (lined bar). The medium for this incubation was Dulbecco's modified Eagle's medium with 4.5 g/L glucose, 1 mM sodium pyruvate, 0.01 mg/ml human transferrin and 10% fetal bovine serum. After 48 hours of incubation the uptake of oleate by enterocytes was measured over a 5 minute time interval. Measurements were done in quadruplicate. The uptake assay was done in Hank's buffered salt solution with 10 mM taurocholate. Only the enterocytes given FATP4 antisense oligonucleotide showed a concentration dependent decrease of fatty acid uptake, inhibiting it at a 100 μM concentration by ~50%. This effect was FATP4 specific, since only the antisense oligonucleotide which can bind to the FATP4 mRNA and block its translation inhibited

uptake, but not a control oligonucleotide differing only in the sequence but not the nucleotide content, ruling out a toxic or otherwise nonspecific inhibitory effect of this oligonucleotide due to its chemical composition.

As a further control experiment, the uptake of oleate was measured along with the uptake of methionine in the same cultured enterocytes. Antisense oligonucleotide, mismatch sequence oligonucleotide, or no oligonucleotide was added to a concentration of 100 μM to cultures of enterocytes. After incubation for 48 hours, the uptake of both ^{3}H-labeled oleate and ^{35}S-labeled methionine was assayed. Results are shown in Figure 42. Fatty acid uptake is at the left side of the paired bars; methionine uptake is on the right side of the paired bars. The fact that amino acid uptake was not influenced by the antisense oligonucleotide treatment further supports the conclusion that the antisense oligonucleotide causes a specific reduction in translation of FATP4-specific mRNA.

Example 14: mmFATP2 Is Expressed in Proximal Renal Tubule Epithelium

Northern analysis showed that mmFATP1, mmFATP2, and mmFATP4 are present in the kidney. In situ hybridization (methods as for Example 6) was performed to determine which cell type(s) of the kidney these mRNAs are expressed in. mmFATP1 mRNA was present in virtually all cells throughout the kidney with no obvious preference for a particular cell type. In contrast, mmFATP2 was expressed only in the renal cortex. Within the cortex, expression of mmFATP2 was restricted to the epithelial cells of the proximal renal tubules. The primary function of proximal renal tubule cells is the reabsorption of filtered salts and nutrients (e.g., glucose), a process that requires mitochondrial oxidation and that can utilize fatty acids as energy substrates. Based on the localization of mmFATP2, it is possible that mmFATP2 is important for reabsorption in the kidney by allowing uptake of an energy source (fatty acids) from the blood into renal epithelial cells. Alternatively, if fatty acids need to be reabsorbed in the kidney, similarly to glucose, FATP2 could be involved in the reabsorption of fatty acids. Determination of the subcellular localization of FATP2 will distinguish between these two possibilities.

Table 5. Mouse FATP mRNA Expression

Mouse Probes	mFATP1	mFATP2	mFATP3	mFATP4	mFATP5
E18.5 embryo expression	everywhere, brain = thymus> heart> brown fat, others	liver (hepatocytes)	-	Brain, small intestine, superior cervical ganglion (SCG), dorsal root ganglion (DRG), other regions have lower expression	Mouse Probes
Duodenum	-	villi (surface epithelium)	-	villi (surface epithelium)	-
Jejunum	-	villi (surface epithelium)	-	villi (surface epithelium)	-
Ileum	-	villi (surface epithelium)	-	villi (surface epithelium)	-
Colon	low expression in the crypt	very low level in the crypt	-	-	-
Kidney	cortex and medulla	proximal tubules	-	-	-

Table 5 (continued). Mouse FATP mRNA Expression

Mouse Probes	mFATP1	mFATP2	mFATP3	mFATP4	mFATP5
Liver	-	hepatocytes	hepatocytes	-	hepatocytes
Pancreas	exocrine secretory units or acinar cells; endocrine pancreas (islet) are negative	exocrine secretory units or acinar cells; endocrine pancreas (islet) are negative	-	-	-
Brain	Neuronal expression throughout the brain including hypothalamus	-	-	Neuronal expression throughout the brain including hypothalamus	-
Heart	myocytes	-	-		
Testis	seminiferous tubules	-	seminiferous tubules		
Lung	bronchiole	-	-		
Adipose	adipocyte	adipocyte	-		

Example 15: Isolation of full-length human FATP3

Full-length clones encoding human FATP3 were identified by searching databases for sequences similar to the murine FATP1-5 coding regions using the BlastX algorithm (Altschul *et al.*, *J. Mol. Biol. 215*: 403-410, 1990). Human clones with similarity to the 5' end of murine FATP sequences were sequenced completely. A clone encoding full-length human FATP3 was obtained from a human bone library constructed in the mammalian expression vector pMET7 (Tartaglia, L.A. *et al.*, *Cell 83*: 1263-1271, 1995). To identify human cDNA clones encoding FATP family members, databases were searched for sequences similar to murine FATP1-5 coding regions. One clone was found to encode the human ortholog of mmFATP3 and was designated hsFATP3. The DNA and predicted protein sequences of hsFATP3 are shown in Figures 94A and 94B. hsFATP5 is predicted to encode a 703 amino acid 75.6 kD protein with multiple membrane-spanning domains. A comparison of the DNA sequences of mouse and human FATP3 shows that the mouse and human orthologs are 81% identical to each other within the coding region. At the amino acid level, hsFATP3 is ~86% identical to mm FATP3 within the coding region. The sequence identities between mouse and human FATP3 are considerably higher than those observed between different FATP family members within one species (~40%) and are present in the N-terminal part of the protein, a region that is poorly conserved between different FATP family members.

Example 16: Substrate Specificity of Fatty Acid Transport in hsFATP-Transfected Clones

Using a mammalian expression vector, we generated 40 stable 239 cell lines expressing hsFATP4 and 20 cell lines transfected with a control plasmid. The ability of the different cell lines to take up FA, as assessed by uptake assays using the fluorescently labeled Bodipy-palmitate, correlated well with their FATP4 expression levels determined by Western blotting (FIG. 95). All 20 vector control clones showed amounts of Bodipy-FA uptake similar to each other and to untransfected 239 cells. In contrast, among the 40 FATP4 transfected clones, a large number (~20) showed an

approximately 2-fold increase in Bodipy-FA uptake compared to any of the vector controls, and three had a 5- to 10-fold increase in Bodipy-FA uptake.

Several of the cell lines with the highest amount of Bodipy-FA uptake as well as isolated primary enterocytes were used to measure the uptake of radiolabeled FAs. Short-term uptake by 293 cells and enterocytes of all FAs tested was linear (FIG. 97). hsFATP4 expression enhanced the rate of palmitate uptake approximately 3 fold over 293 cells transfected with vector alone (FIG. 97) and also accelerated the uptake of oleate but not of linolate, arachidonate, octanoate, butyrate or cholesterol (Table 6). Isolated primary enterocytes showed a similar preference for palmitate and oleate, and absence of transport of arachidonate, octanoate, and butyrate, but displayed a more robust transport of linolate and cholesterol than the transfected 293 cells.

To further characterize the substrate specificity of FATP4, we measured the uptake by stably transfected 293 cells of 5 μM Bodipy-FA in the presence of a 20 fold molar excess (i.e., 100 μM) of FAs, FA-derivatives and lipid soluble vitamins and hormones. Both saturated and non-saturated fatty acids containing 10 to 26 C atoms strongly competed for uptake of Bodipy-palmitate (FIG. 96 and Table 7) and thus are presumed to be substrates of FATP4. In contrast, fatty acids with eight of fewer C atoms did not compete and thus are presumed not to be FATP4 substrates. Similarly, esters of long chain FAs and other hydrophobic molecules tested had no effect on uptake of Bodipy-palmitate.

LCFA Uptake Assays (Methods)

Bodipy-FA uptake assays using FACS were performed, adapted to a 96-well format. LCFA uptake assays with enterocytes or with stably transfected 293 cells were done as follows. Mixed micelles of radiolabeled FA (NEN) and taurocholate (Sigma) in HBS were generated by brief sonication at 37°C. Equal volumes of cells and micelle solution were mixed, resulting in a final FA concentration of 25 μM for antisense assays and 10 μM for substrate specificity assays. Final taurocholate concentration was 5 mM. Cells were incubated for the indicated amount of time at 37°C. The assay was stopped by transferring the cells onto filter paper followed by extensive washes with

ice-cold HBS containing 0.1% BSA using a cell harvester (Brandell). Incorporated oleate was then determined by β-scintillation counting (Beckman).

Table 6

Uptake of Different Substrates by FATP4 Expressing Cell Lines and Enterocytes

Fatty Acid	293 Cells Control*	293 Cells Stably Expressing FATP4	FATP4 specific	Enterocytes*
Palmitate	564	1695	1131	3036
Oleate	662	1122	459	117
Linolate	640	673	33	116
Arachidonate	3	5	2	0
Octanoate	0	0	0	5
Butyrate	0	50	50	73
Cholesterol	319	345	26	531

Uptake of different substrates by enterocytes and by control and stable FATP4-expressing 293 cells. The rates of uptake for the indicated fatty acids was measured over 4 min taking measurements every 30 s. All fatty acids were at a concentration of 10 μM in HBS containing 5 mM taurocholate.

*Uptake measured as *pmol/min 10^6 cells*

Table 7

Competition of Bodipy-FA Uptake by FATP4 Expressing Cells

Fatty Acids	Formula	Competition
Butyric Acid	$C_4H_8O_2$	-
Caproic Acid	$C_6H_{12}O_2$	-
Caprylic Acid	$C_8H_{16}O_2$	-
Capric Acid	$C_{10}H_{20}O_2$	++
Lauric Acid	$C_{12}H_{24}O_2$	++
Myristic Acid	$C_{14}H_{28}O_2$	++
Palmitic Acid	$C_{16}H_{32}O_2$	++
Stearic Acid	$C_{18}H_{36}O_2$	+
Oleic Acid	$C_{18}H_{34}O_2$	++
Linoleic Acid	$C_{18}H_{32}O_2$	++
Arachidic Acid	$C_{20}H_{40}O_2$	++
Lignoceric Acid	$C_{24}H_{48}O_2$	++
Cerotic Acid	$C_{26}H_{52}O_2$	++

Fatty Acid Derivatives

Fatty Acids	Formula	Competition
Palmitic Acid Methyl Ester	$C_{17}H_{34}O_2$	-
Stearic Acid Methyl Ester	$C_{19}H_{38}O_2$	-
Oleic Acid Ethyl Ester	$C_{20}H_{38}O_2$	-
Oleic Acid Oley Ester	$C_{36}H_{68}O_2$	-
Oleoyl CoA	$C_{39}H_{68}N_7O_{17}P_3S$	-
Cholesteryl Oleate	$C_{45}H_{78}O_2$	-

Table 7 Continued

Competition of Bodipy-FA Uptake by FATP4 Expressing Cells

Lipid-Soluble Vitamins & Homones

Fatty Acids	Formula	Competition
Retinoic Acid (Pro-Vitamin A)	$C_{20}H_{28}O_2$	±
Ergocalciferol (Vitamin D2)	$C_{28}H_{44}O_2$	-
Tocopherol (Vitamin E)	$C_{29}H_{50}O_2$	-
3-Phytylamenadione (Vitamin K1)	$C_{31}H_{46}O_2$	-
Prostaglandin E2	$C_{20}H_{32}O_5$	-

Competition for Bodipy-FA uptake by FATP4 expressing cells by different hydrophobic compounds. The uptake of 5 μM Bodipy-FA, C1-Bodipy-C12 was measured in the presence of a 20-fold molar excess (i.e., 100 μM) of the indicated fatty acids or fatty acid derivatives. The maximal 100% inhibition was defined as the amount of Bodipy-FA incorporated in the presence of 200 μM lauric acid which was on average 18% ± 5% that of untreated cells.

-: 0% - 30% inhibition by the indicated substance

±: 30% - 50% inhibition

+: 50% - 70% inhibition

++: 70% - 100% inhibition

All references cited herein are incorporated by reference in their entirety.

While this invention has been particularly shown and described with references to preferred embodiments thereof, it will be understood by those skilled in the art that various changes in form and details may be made therein without departing from the spirit and scope of the invention as defined by the appended claims.

CLAIMS

What is claimed is:

1. A method for identifying an agent which binds to a protein comprising an amino acid sequence of SEQ ID NO:49 or SEQ ID NO:53, comprising the steps of contacting the agent with the isolated protein under conditions appropriate for binding of the agent to the isolated protein, and detecting a resulting agent-protein complex.

2. The method of Claim 1 wherein the step of contacting the agent with isolated protein is performed in an artificial membrane system.

3. The method of Claim 1 wherein the isolated protein is in isolated plasma membrane.

4. A method for identifying an agent which inhibits interaction between an isolated protein comprising amino acid sequence SEQ ID NO:49, or SEQ ID NO:53, and further comprising a ligand of said protein, comprising:

(a) combining:

(1) said isolated protein;

(2) the ligand of said protein; and

(3) a candidate agent to be assessed for its ability to inhibit interaction between said protein of (1) and the ligand of (2), under conditions appropriate for interaction between the said protein of (1) and the ligand of (2);

(b) determining the extent to which said protein of (1) and the ligand of (2) interact; and

(c) comparing the extent determined in (b) with the extent to which interaction of said protein of (1) and the ligand of (2) occurs in the

absence of the candidate agent to be assessed and under the same conditions appropriate for interaction of said protein of (1) with the ligand of (2);

wherein if the extent to which interaction of said protein of (1) and the ligand of (2) occurs is less in the presence of the candidate agent than in the absence of the candidate agent, the candidate agent is an agent which inhibits interaction between said protein and the ligand of said protein.

5. The method of Claim 4 wherein (a) is performed in an artificial membrane system.

6. The method of Claim 4 wherein said isolated protein is in isolated plasma membrane.

7. A method for identifying an agent which binds to a protein, said protein encoded by (1) a polynucleotide comprising a nucleotide sequence which encodes a naturally occurring allelic variant of a polypeptide consisting of the amino acid sequence of FATP2, wherein said polynucleotide hybridizes to a complement of a polynucleotide consisting of SEQ ID NO:48 under high stringency conditions, or by (2) a polynucleotide comprising a nucleotide sequence which encodes a naturally occurring allelic variant of a polypeptide consisting of the amino acid sequence of FATP4, wherein said polynucleotide hybridizes to a complement of a polynucleotide consisting of SEQ ID NO:52 under high stringency conditions, comprising the steps of isolating the protein, contacting the agent with the isolated protein under conditions appropriate for binding of the agent to the isolated protein, and detecting a resulting agent-protein complex.

8. The method of Claim 7 wherein the step of contacting the agent with the isolated protein is performed in an artificial membrane system.

9. The method of Claim 7 wherein the isolated protein is in isolated plasma membrane.

10. A method for identifying an agent which inhibits interaction between (1) an isolated protein, said protein being encoded by (i) a polynucleotide comprising a nucleotide sequence which encodes a naturally occurring allelic variant of a polypeptide consisting of the amino acid sequence of FATP2, wherein said polynucleotide hybridizes to a complement of a polynucleotide consisting of SEQ ID NO:48 under high stringency conditions, or by (ii) a polynucleotide having a nucleotide sequence which encodes a naturally occurring allelic variant of a polypeptide consiting of the amino acid sequence of FATP4, wherein said polynucleotide hybridizes to a complement of a polynucleotide consisting of SEQ ID NO:52 under high stringency conditions and (2) a ligand of said protein, comprising:

(a) combining:

(1) said isolated protein;

(2) the ligand of said protein; and

(3) a candidate agent to be assessed for its ability to inhibit interaction between said protein of (1) and the ligand of (2), under conditions appropriate for interaction between said protein of (1) and the ligand of (2);

(b) determining the extent to which said protein of (1) and the ligand of (2) interact; and

(c) comparing the extent determined in (b) with the extent to which interaction of said protein of (1) and the ligand of (2) occurs in the absence of the candidate agent to be assessed and under the same conditions appropriate for interaction of said protein of (1) with the ligand of (2);

wherein if the extent to which interaction of said protein of (1) and the ligand of (2) occurs is less in the presence of the candidate agent than in the absence of the

candidate agent, the candidate agent is an agent which inhibits interaction between said protein and the ligand of said protein.

11. The method of Claim 10 wherein (a) is performed in an artificial membrane system.

12. The method of Claim 10 wherein said isolated protein is in isolated plasma membrane.

13. A method for identifying an agent which binds to a protein encoded by a nucleic acid encoding a fatty acid transport protein comprising an amino acid sequence sharing at least about 95% amino acid sequence similarity with the amino acid sequence in SEQ ID NO:49, or SEQ ID NO:53 comprising the steps of isolating the protein, contacting the agent with the isolated protein under conditions appropriate for binding of the agent to the isolated protein, and detecting a resulting agent-protein complex.

14. The method of Claim 13 wherein the step of contacting the agent with isolated protein is performed in an artificial membrane system.

15. The method of Claim 13 wherein the isolated protein is in isolated plasma membrane.

16. A method for identifying an agent which inhibits interaction between (i) an isolated protein encoded by a nucleic acid encoding a fatty acid transport protein comprising an amino acid sequence sharing at least about 90% amino acid sequence similarity with the amino acid sequence in SEQ ID NO:49, or (ii) a protein encoded by a nucleic acid encoding a fatty acid transport protein comprising an amino acid sequence sharing at least about 90% amino acid

sequence similarity with the amino acid sequence in SEQ ID NO:53 and a ligand of said protein, said method comprising:

(a) combining:

(1) said isolated protein;

(2) the ligand of said protein; and

(3) a candidate agent to be assessed for its ability to inhibit interaction between said protein of (1) and the ligand of (2), under conditions appropriate for interaction between the said protein of (1) and the ligand of (2);

(b) determining the extent to which said protein of (1) and the ligand of (2) interact; and

(c) comparing the extent determined in (b) with the extent to which interaction of said protein of (1) and the ligand of (2) occurs in the absence of the candidate agent to be assessed and under the same conditions appropriate for interaction of said protein of (1) with the ligand of (2);

wherein if the extent to which interaction of said protein of (1) and the ligand of (2) occurs is less in the presence of the candidate agent than in the absence of the candidate agent, the candidate agent is an agent which inhibits interaction between said protein and the ligand of said protein.

17. The method of Claim 16 wherein (a) is performed in an artificial membrane system.

18. The method of Claim 16 wherein said isolated protein is in isolated plasma membrane.

19. A method for identifying an agent which is an inhibitor of fatty acid uptake by (i) a protein encoded by a polynucleotide comprising a nucleotide sequence which encodes a protein consisting of the amino acid sequence in SEQ ID

NO:49, or by (ii) a protein encoded by a polynucleotide comprising a nucleotide sequence which encodes a protein consisting of the amino acid sequence in SEQ ID NO:53, comprising the steps of:

a) maintaining test cells expressing said polynucleotide in the presence of a fatty acid and an agent to be tested as an inhibitor of fatty acid uptake;

b) measuring uptake of the fatty acid in the test cells; and

c) comparing uptake of the fatty acid in the test cells with uptake of the fatty acid in suitable control cells;

wherein lower uptake of the fatty acid in the test cells compared to uptake of the fatty acid in the control cells is indicative that the agent is an inhibitor of fatty acid uptake by said protein.

20. An inhibitor of fatty acid uptake identified by the method of Claim 19.

21. The method of Claim 19 further comprising the steps of:

a) administering the agent to one or more test animals;

b) measuring exogenously supplied fatty acids in one or more samples of tissue or bodily fluid from said test animals;

c) measuring exogenously supplied fatty acids in one or more comparable samples of tissue or bodily fluid from suitable control animals;

d) comparing the fatty acids of b) with the fatty acids of c);

whereby, lower fatty acids in step b) than in step c) is indicative that the agent is an inhibitor of said protein.

22. An inhibitor of fatty acid uptake identified by the method of Claim 21.

23. A method for identifying an agent which is an inhibitor of fatty acid uptake by a protein, said protein encoded by (i) a polynucleotide comprising a nucleotide sequence which encodes a naturally occurring allelic variant of a polypeptide consisting of the amino acid sequence of FATP2, wherein said polynucleotide

hybridizes to a complement of a polynucleotide consisting of SEQ ID NO:48 under high stringency conditions, or by (ii) a polynucleotide comprising a nucleotide sequence which encodes a naturally occurring allelic variant of a polypeptide consisting of the amino acid sequence of FATP4, wherein said polynucleotide hybridizes to a complement of a polynucleotide consisting of SEQ ID NO:52 under high stringency conditions, comprising the steps of:

a) maintaining test cells expressing said polynucleotide in the presence of a fatty acid and an agent to be tested as an inhibitor of fatty acid uptake;

b) measuring uptake of the fatty acid in the test cells; and

c) comparing uptake of the fatty acid in the test cells with uptake of the fatty acid in suitable control cells;

wherein lower uptake of the fatty acid in the test cells compared to uptake of the fatty acid in the control cells is indicative that the agent is an inhibitor of fatty acid uptake by said protein.

24. An inhibitor of fatty acid uptake identified by the method of Claim 23.

25. The method of Claim 23 further comprising the steps of:

a) administering the agent to one or more test animals;

b) measuring exogenously supplied fatty acids in one or more samples of tissue or bodily fluid from said test animals;

c) measuring exogenously supplied fatty acids in one or more comparable samples of tissue or bodily fluid from suitable control animals;

d) comparing the fatty acids of b) with the fatty acids of c);

whereby, lower fatty acids in step b) than in step c) is indicative that the agent is an inhibitor of said protein.

26. An inhibitor of fatty acid uptake identified by the method of Claim 25.

27. A method for identifying an agent which is an inhibitor of fatty acid uptake by a protein, said protein being encoded by (i) a nucleic acid encoding a fatty acid transport protein comprising an amino acid sequence sharing at least about 95% amino acid sequence similarity with the amino acid sequence in SEQ ID NO:49 or by (ii) a nucleic acid encoding a fatty acid transport protein comprising an amino acid sequence sharing at least about 95% amino acid sequence similarity with the amino acid sequence in SEQ ID NO:53, comprising the steps of:
    a) maintaining test cells expressing said polynucleotide in the presence of a fatty acid and an agent to be tested as an inhibitor of fatty acid uptake;
    b) measuring uptake of the fatty acid in the test cells; and
    c) comparing uptake of the fatty acid in the test cells with uptake of the fatty acid in suitable control cells;

    wherein lower uptake of the fatty acid in the test cells compared to uptake of the fatty acid in the control cells is indicative that the agent is an inhibitor of fatty acid uptake by said protein.

28. An inhibitor of fatty acid uptake identified by the method of Claim 27.

29. The method of Claim 27 further comprising the steps of:
    a) administering the agent to one or more test animals;
    b) measuring exogenously supplied fatty acids in one or more samples of tissue or bodily fluid from said test animals;
    c) measuring exogenously supplied fatty acids in one or more comparable samples of tissue or bodily fluid from suitable control animals;
    d) comparing the fatty acids of b) with the fatty acids of c);

    whereby, lower fatty acids in step b) than in step c) is indicative that the agent is an inhibitor of said protein.

30. An inhibitor of fatty acid uptake identified by the method of Claim 27.

31. A method for identifying an agent which is an inhibitor of (i) a protein encoded by a polynucleotide comprising a nucleotide sequence which encodes a protein comprising the amino acid sequence in SEQ ID NO:49 or (ii) a protein encoded by a polynucleotide comprising a nucleotide sequence which encodes a protein comprising the amino acid sequence in SEQ ID NO:53, comprising the steps of:

(a) introducing into host cells one or more vectors comprising a polynucleotide expressing said protein;

(b) culturing a first aliquot of the host cells with fatty acid substrate of said protein and with an agent being tested as an inhibitor of said protein;

(c) culturing a second aliquot of the host cells with fatty acid substrate of said protein;

(d) measuring, in the first and second aliquots, uptake of the fatty acid substrate of the host cells;

wherein less uptake of the fatty acid substrate in the first aliquot compared to the second aliquot is indicative that the agent is an inhibitor of said protein.

32. The method of Claim 31 further comprising the steps of:

a) administering the agent to one or more test animals;

b) measuring exogenously supplied fatty acids in one or more samples of tissue or bodily fluid from suitable control animals;

c) measuring exogenously supplied fatty acids in one or more comparable samples of tissue or bodily fluid from suitable control animals; and

d) comparing the fatty acids of b) with the fatty acids of c);

whereby, lower fatty acids in step b) than in step c) is indicative that the agent is an inhibitor of said protein.

33. A method for identifying an agent which is an inhibitor of a protein, said protein being encoded by (i) a polynucleotide comprising a nucleotide sequence which encodes a naturally occurring allelic variant of a polypeptide consisting of the amino acid sequence of FATP2, wherein said polynucleotide hybridizes to a

complement of a polynucleotide consisting of SEQ ID NO:48 under high stringency conditions, or by (ii) a polynucleotide comprising a nucleotide sequence which encodes a naturally occurring allelic variant of a polypeptide consisting of the amino acid sequence of FATP4, wherein said polynucleotide hybridizes to a complement of a polynucleotide consisting of SEQ ID NO:52 under high stringency conditions, comprising the steps of:

(a) introducing into host cells one or more vectors comprising a polynucleotide expressing said protein;

(b) culturing a first aliquot of the host cells with fatty acid substrate of said protein and with an agent being tested as an inhibitor of said protein;

(c) culturing a second aliquot of the host cells with fatty acid substrate of said protein;

(d) measuring, in the first and second aliquots, uptake of the fatty acid substrate of the host cells;

wherein less uptake of the fatty acid substrate in the first aliquot compared to the second aliquot is indicative that the agent is an inhibitor of said protein.

34. The method of Claim 33 further comprising the steps of:

a) administering the agent to one or more test animals;

b) measuring exogenously supplied fatty acids in one or more samples of tissue or bodily fluid from suitable control animals;

c) measuring exogenously supplied fatty acids in one or more comparable samples of tissue or bodily fluid from suitable control animals; and

d) comparing the fatty acids of b) with the fatty acids of c);

whereby, lower fatty acids in step b) than in step c) is indicative that the agent is an inhibitor of said protein.

35. A method for identifying an agent which is an inhibitor of a protein, said protein being encoded by (i) a nucleic acid encoding a fatty acid transport protein comprising an amino acid sequence sharing at least about 95% amino acid

sequence similarity with the amino acid sequence in SEQ ID NO:49 or by (ii) a nucleic acid encoding a fatty acid transport protein comprising an amino acid sequence sharing at least about 95% amino acid sequence similarity with the amino acid sequence in SEQ ID NO:53, comprising the steps of:

(a) introducing into host cells one or more vectors comprising a polynucleotide expressing said protein;

(b) culturing a first aliquot of the host cells with fatty acid substrate of said protein and with an agent being tested as an inhibitor of said protein;

(c) culturing a second aliquot of the host cells with fatty acid substrate of said protein;

(d) measuring, in the first and second aliquots, uptake of the fatty acid substrate of the host cells;

wherein less uptake of the fatty acid substrate in the first aliquot compared to the second aliquot is indicative that the agent is an inhibitor of said protein.

36. The method of Claim 35 further comprising the steps of:

a) administering the agent to one or more test animals;

b) measuring exogenously supplied fatty acids in one or more samples of tissue or bodily fluid from suitable control animals;

c) measuring exogenously supplied fatty acids in one or more comparable samples of tissue or bodily fluid from suitable control animals; and

d) comparing the fatty acids of b) with the fatty acids of c).

whereby, lower fatty acids in step b) than in step c) is indicative that the agent is an inhibitor of said protein.

37. A method for identifying an agent which binds to a protein comprising an amino acid sequence of SEQ ID NO:57, comprising the steps of contacting the agent with the isolated protein under conditions appropriate for binding of the agent to the isolated protein, and detecting a resulting agent-protein complex.

38. The method of Claim 37 wherein the step of contacting the agent with isolated protein is performed in an artificial membrane system.

39. The method of Claim 37 wherein the isolated protein is in isolated plasma membrane.

40. A method for identifying an agent which inhibits interaction between an isolated protein comprising an amino acid sequence of SEQ ID NO:57, and further comprising a ligand of said protein, comprising:

(a) combining:

(1) said isolated protein;

(2) the ligand of said protein; and

(3) a candidate agent to be assessed for its ability to inhibit interaction between said protein of (1) and the ligand of (2), under conditions appropriate for interaction between the said protein of (1) and the ligand of (2);

(b) determining the extent to which said protein of (1) and the ligand of (2) interact; and

(c) comparing the extent determined in (b) with the extent to which interaction of said protein of (1) and the ligand of (2) occurs in the absence of the candidate agent to be assessed and under the same conditions appropriate for interaction of said protein of (1) with the ligand of (2);

wherein if the extent to which interaction of said protein of (1) and the ligand of (2) occurs is less in the presence of the candidate agent than in the absence of the candidate agent, the candidate agent is an agent which inhibits interaction between said protein and the ligand of said protein.

41. The method of Claim 40 wherein (a) is performed in an artificial membrane system.

42. The method of Claim 40 wherein said isolated protein is in isolated plasma membrane.

43. A method for identifying an agent which binds to a protein, said protein encoded by a polynucleotide comprising a nucleotide sequence which encodes a naturally occurring allelic variant of a polypeptide consisting of the amino acid sequence of FATP6, wherein said polynucleotide hybridizes to a complement of a polynucleotide consisting of SEQ ID NO:56 under high stringency conditions, comprising the steps of isolating the protein, contacting the agent with the isolated protein under conditions appropriate for binding of the agent to the isolated protein, and detecting a resulting agent-protein complex.

44. The method of Claim 43 wherein the step of contacting the agent with the isolated protein is performed in an artificial membrane system.

45. The method of Claim 43 wherein the isolated protein is in isolated plasma membrane.

46. A method for identifying an agent which inhibits interaction between (1) an isolated protein, said protein encoded by a polynucleotide comprising a nucleotide sequence which encodes a naturally occurring allelic variant of a polypeptide consisting of the amino acid sequence of FATP6, wherein said polynucleotide hybridizes to a complement of a polynucleotide consisting of SEQ ID NO:56 under high stringency conditions, and (2) a ligand of said protein, comprising:

(a) combining:

(1) said isolated protein;

(2) the ligand of said protein; and

(3) a candidate agent to be assessed for its ability to inhibit interaction between said protein of (1) and the ligand of (2),

under conditions appropriate for interaction between said protein of (1) and the ligand of (2);

(b) determining the extent to which said protein of (1) and the ligand of (2) interact; and

(c) comparing the extent determined in (b) with the extent to which interaction of said protein of (1) and the ligand of (2) occurs in the absence of the candidate agent to be assessed and under the same conditions appropriate for interaction of said protein of (1) with the ligand of (2);

wherein if the extent to which interaction of said protein of (1) and the ligand of (2) occurs is less in the presence of the candidate agent than in the absence of the candidate agent, the candidate agent is an agent which inhibits interaction between said protein and the ligand of said protein.

47. The method of Claim 46 wherein (a) is performed in an artificial membrane system.

48. The method of Claim 46 wherein the isolated protein is in isolated plasma membrane.

49. A method for identifying an agent which binds to a protein encoded by a nucleic acid encoding a fatty acid transport protein consisting of an amino acid sequence sharing at least about 95% amino acid sequence similarity with the amino acid sequence in SEQ ID NO:57 comprising the steps of isolating the protein, contacting the agent with the isolated protein under conditions appropriate for binding of the agent to the isolated protein, and detecting a resulting agent-protein complex.

50. The method of Claim 49 wherein the step of contacting the agent with isolated protein is performed in an artificial membrane system.

51. The method of Claim 49 wherein the isolated protein is in isolated plasma membrane.

52. A method for identifying an agent which inhibits interaction between an isolated protein encoded by a nucleic acid encoding a fatty acid transport protein comprising an amino acid sequence sharing at least about 90% amino acid sequence similarity with the amino acid sequence in SEQ ID NO:57 and a ligand of said protein, said method comprising:

(a) combining:

(1) said isolated protein;

(2) the ligand of said protein; and

(3) a candidate agent to be assessed for its ability to inhibit interaction between said protein of (1) and the ligand of (2), under conditions appropriate for interaction between the said protein of (1) and the ligand of (2);

(b) determining the extent to which said protein of (1) and the ligand of (2) interact; and

(c) comparing the extent determined in (b) with the extent to which interaction of said protein of (1) and the ligand of (2) occurs in the absence of the candidate agent to be assessed and under the same conditions appropriate for interaction of said protein of (1) with the ligand of (2);

wherein if the extent to which interaction of said protein of (1) and the ligand of (2) occurs is less in the presence of the candidate agent than in the absence of the candidate agent, the candidate agent is an agent which inhibits interaction between said protein and the ligand of said protein.

53. The method of Claim 52 wherein (a) is performed in an artificial membrane system.

54. The method of Claim 52 wherein said isolated protein is in isolated plasma membrane.

55. A method for identifying an agent which is an inhibitor of fatty acid uptake by a protein encoded by a polynucleotide comprising a nucleotide sequence which encodes a protein consisting of the amino acid sequence in SEQ ID NO:57, comprising the steps of:

a) maintaining test cells expressing said polynucleotide in the presence of a fatty acid and an agent to be tested as an inhibitor of fatty acid uptake;

b) measuring uptake of the fatty acid in the test cells; and

c) comparing uptake of the fatty acid in the test cells with uptake of the fatty acid in suitable control cells;

wherein lower uptake of the fatty acid in the test cells compared to uptake of the fatty acid in the control cells is indicative that the agent is an inhibitor of fatty acid uptake by said protein.

56. An inhibitor of fatty acid uptake identified by the method of Claim 55.

57. The method of Claim 55 further comprising the steps of:

a) administering the agent to one or more test animals;

b) measuring exogenously supplied fatty acids in one or more samples of tissue or bodily fluid from said test animals;

c) measuring exogenously supplied fatty acids in one or more comparable samples of tissue or bodily fluid from suitable control animals;

d) comparing the fatty acids of b) with the fatty acids of c);

whereby, lower fatty acids in step b) than in step c) is indicative that the agent is an inhibitor of said protein.

58. An inhibitor of fatty acid uptake identified by the method of Claim 57.

59. A method for identifying an agent which is an inhibitor of fatty acid uptake by a protein, said protein encoded by a polynucleotide comprising a nucleotide sequence which encodes a naturally occurring allelic variant of a polypeptide consisting of the amino acid sequence of FATP6, wherein said polynucleotide hybridizes to a complement of a polynucleotide consisting of SEQ ID NO:56 under high stringency conditions, comprising the steps of:

a) maintaining test cells expressing said polynucleotide in the presence of a fatty acid and an agent to be tested as an inhibitor of fatty acid uptake;

b) measuring uptake of the fatty acid in the test cells; and

c) comparing uptake of the fatty acid in the test cells with uptake of the fatty acid in suitable control cells;

wherein lower uptake of the fatty acid in the test cells compared to uptake of the fatty acid in the control cells is indicative that the agent is an inhibitor of fatty acid uptake by said protein.

60. An inhibitor of fatty acid uptake identified by the method of Claim 59.

61. The method of Claim 59 further comprising the steps of:

a) administering the agent to one or more test animals;

b) measuring exogenously supplied fatty acids in one or more samples of tissue or bodily fluid from said test animals;

c) measuring exogenously supplied fatty acids in one or more comparable samples of tissue or bodily fluid from suitable control animals;

d) comparing the fatty acids of b) with the fatty acids of c);

whereby, lower fatty acids in step b) than in step c) is indicative that the agent is an inhibitor of said protein.

62. An inhibitor of fatty acid uptake identified by the method of Claim 61.

63. A method for identifying an agent which is an inhibitor of fatty acid uptake by a protein, said protein being encoded by a nucleic acid encoding a fatty acid transport protein comprising an amino acid sequence sharing at least about 95% amino acid sequence similarity with the amino acid sequence in SEQ ID NO:57, comprising the steps of:

a) maintaining test cells expressing said polynucleotide in the presence of a fatty acid and an agent to be tested as an inhibitor of fatty acid uptake;

b) measuring uptake of the fatty acid in the test cells; and

c) comparing uptake of the fatty acid in the test cells with uptake of the fatty acid in suitable control cells;

wherein lower uptake of the fatty acid in the test cells compared to uptake of the fatty acid in the control cells is indicative that the agent is an inhibitor of fatty acid uptake by said protein.

64. An inhibitor of fatty acid uptake identified by the method of Claim 63.

65. The method of Claim 63 further comprising the steps of:

a) administering the agent to one or more test animals;

b) measuring exogenously supplied fatty acids in one or more samples of tissue or bodily fluid from said test animals;

c) measuring exogenously supplied fatty acids in one or more comparable samples of tissue or bodily fluid from suitable control animals;

d) comparing the fatty acids of b) with the fatty acids of c);

whereby, lower fatty acids in step b) than in step c) is indicative that the agent is an inhibitor of said protein.

66. An inhibitor of fatty acid uptake identified by the method of Claim 65.

67. A method for identifying an agent which is an inhibitor of a protein encoded by a polynucleotide comprising a nucleotide sequence which encodes a protein comprising the amino acid sequence in SEQ ID NO:57, comprising the steps of:

(a) introducing into host cells one or more vectors comprising a polynucleotide expressing said protein;

(b) culturing a first aliquot of the host cells with fatty acid substrate of said protein and with an agent being tested as an inhibitor of said protein;

(c) culturing a second aliquot of the host cells with fatty acid substrate of said protein;

(d) measuring, in the first and second aliquots, uptake of the fatty acid substrate of the host cells;

wherein less uptake of the fatty acid substrate in the first aliquot compared to the second aliquot is indicative that the agent is an inhibitor of said protein.

68. The method of Claim 67 further comprising the steps of:

a) administering the agent to one or more test animals;

b) measuring exogenously supplied fatty acids in one or more samples of tissue or bodily fluid from suitable control animals;

c) measuring exogenously supplied fatty acids in one or more comparable samples of tissue or bodily fluid from suitable control animals; and

d) comparing the fatty acids of b) with the fatty acids of c);

whereby, lower fatty acids in step b) than in step c) is indicative that the agent is an inhibitor of said protein.

69. A method for identifying an agent which is an inhibitor of a protein, said protein being encoded by a polynucleotide comprising a nucleotide sequence which encodes a naturally occurring allelic variant of a polypeptide consisting of the amino acid sequence of FATP6, wherein said polynucleotide hybridizes to a complement of a polynucleotide consisting of SEQ ID NO:56 under high stringency conditions, comprising the steps of:

(a) introducing into host cells one or more vectors comprising a polynucleotide expressing said protein;

(b) culturing a first aliquot of the host cells with fatty acid substrate of said protein and with an agent being tested as an inhibitor of said protein;

(c) culturing a second aliquot of the host cells with fatty acid substrate of said protein;

(d) measuring, in the first and second aliquots, uptake of the fatty acid substrate of the host cells;

wherein less uptake of the fatty acid substrate in the first aliquot compared to the second aliquot is indicative that the agent is an inhibitor of said protein.

70. The method of Claim 69 further comprising the steps of:

a) administering the agent to one or more test animals;

b) measuring exogenously supplied fatty acids in one or more samples of tissue or bodily fluid from suitable control animals;

c) measuring exogenously supplied fatty acids in one or more comparable samples of tissue or bodily fluid from suitable control animals; and

d) comparing the fatty acids of b) with the fatty acids of c);

whereby, lower fatty acids in step b) than in step c) is indicative that the agent is an inhibitor of said protein.

71. A method for identifying an agent which is an inhibitor of a protein, said protein being encoded by a nucleic acid encoding a fatty acid transport protein comprising an amino acid sequence sharing at least about 95% amino acid sequence homology with the amino acid sequence in SEQ ID NO:57, comprising the steps of:

(a) introducing into host cells one or more vectors comprising a polynucleotide expressing said protein;

(b) culturing a first aliquot of the host cells with fatty acid substrate of said protein and with an agent being tested as an inhibitor of said protein;

(c) culturing a second aliquot of the host cells with fatty acid substrate of said protein;

(d) measuring, in the first and second aliquots, uptake of the fatty acid substrate of the host cells;

wherein less uptake of the fatty acid substrate in the first aliquot compared to the second aliquot is indicative that the agent is an inhibitor of said protein.

72. The method of Claim 71 further comprising the steps of:

a) administering the agent to one or more test animals;

b) measuring exogenously supplied fatty acids in one or more samples of tissue or bodily fluid from suitable control animals;

c) measuring exogenously supplied fatty acids in one or more comparable samples of tissue or bodily fluid from suitable control animals; and

d) comparing the fatty acids of b) with the fatty acids of c).

whereby, lower fatty acids in step b) than in step c) is indicative that the agent is an inhibitor of said protein.

73. A method for identifying an agent which is an inhibitor of a fatty acid transport protein, comprising the steps of:

(a) introducing into cells one or more vectors comprising a gene encoding a cell surface protein and a nucleic acid encoding the fatty acid transport protein;

(b) contacting the host cells with anti-cell surface protein antibody and labeled fatty acid substrate of the fatty acid transport protein;

(c) contacting a first aliquot of the host cells with an agent being tested as an inhibitor of the fatty acid transport protein, while leaving a second aliquot of the host cells uncontacted with the agent;

(d) identifying, in the first and second aliquots, the host cells expressing the cell surface protein by detecting the anti-cell surface protein antibody bound to the host cells; and

(e) measuring, in the first and second aliquots, uptake of the fatty acid substrate of the host cells identified as expressing the cell surface protein;

wherein less uptake of the fatty acid substrate in the first aliquot compared to the second aliquot is indicative that the agent is an inhibitor of the fatty acid transport protein.

74. The method of Claim 73 wherein the host cells regulably express the FATP4 gene.

75. The method of Claim 73 wherein the host cells are prokaryotes.

76. The method of Claim 73 wherein the prokaryotes are *E. coli*.

77. The method of Claim 73 wherein the fatty acid is a radioactively labeled fatty acid.

78. A method for identifying an agent which is an inhibitor of FATP4, comprising the steps of:

(a) introducing into cells one or more vectors comprising a gene encoding a cell surface protein and a nucleic acid encoding FATP4;

(b) contacting the host cells with anti-cell surface protein antibody and labeled fatty acid substrate of FATP4;

(c) contacting a first aliquot of the host cells with an agent being tested as an inhibitor of FATP4, while leaving a second aliquot of the host cells uncontacted with the agent;

(d) identifying, in the first and second aliquots, the host cells expressing the cell surface protein by detecting the anti-cell surface protein antibody bound to the host cells; and

(e) measuring, in the first and second aliquots, uptake of the fatty acid substrate of the host cells identified as expressing the cell surface protein;

wherein less uptake of the fatty acid substrate in the first aliquot compared to the second aliquot is indicative that the agent is an inhibitor of FATP4.

79. The method of Claim 78 further comprising the steps of:

a) administering the agent to one or more test animals;

b) measuring exogenously supplied fatty acids in one or more samples of tissue or bodily fluid from suitable control animals;

c) measuring exogenously supplied fatty acids in one or more comparable samples of tissue or bodily fluid from suitable control animals; and

d) comparing the fatty acids of b) with the fatty acids of c);

whereby, lower fatty acids in step b) than in step c) is indicative that the agent is an inhibitor of said protein.

80. The method of Claim 78 wherein the cell surface protein is CD2.

81. The method of Claim 78 wherein the fatty acid substrate is BODIPY-labeled.

FATTY ACID TRANSPORT PROTEINS

ABSTRACT OF THE DISCLOSURE

A family of fatty acid transport proteins (FATPs) mediate transport of long chain fatty acids (LCFAs) across cell membranes into cells. These proteins exhibit different expression patterns among the organs of mammals. Nucleic acids encoding FATPs of this family, vectors comprising these nucleic acids, as well as the production of FATP proteins in host cells are described. Also described are methods to test FATPs for fatty acid transport function, and methods to identify inhibitors or enhancers of transport function. The altering of LCFA uptake by administering to the mammal an inhibitor or enhancer of FATP transport function of a FATP in the small intestine can decrease or increase calories available as fats, and can decrease or increase circulating fatty acids. The organ specificity of FATP distribution can be exploited in methods to direct drugs, diagnostic indicators and so forth to an organ such as the heart.

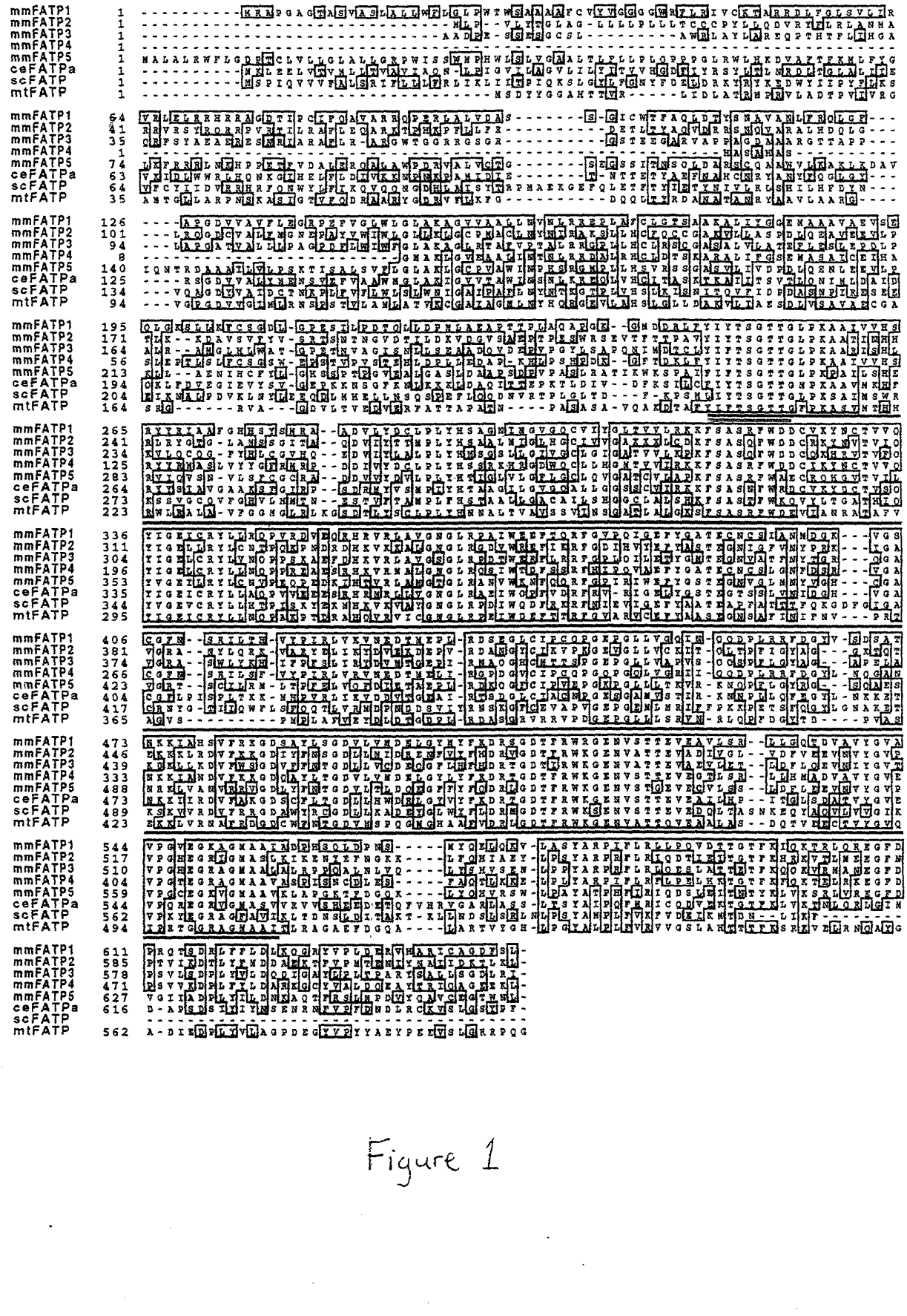

mmFATP1
mmFATP2
mmFATP3
mmFATP4
mmFATP5
ceFATPa
scFATP
mtFATP

Figure 1

Fig 2A

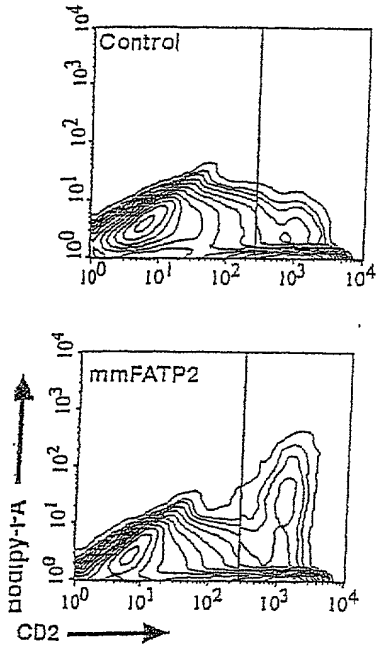


Fig. 2C

Fig. 2B

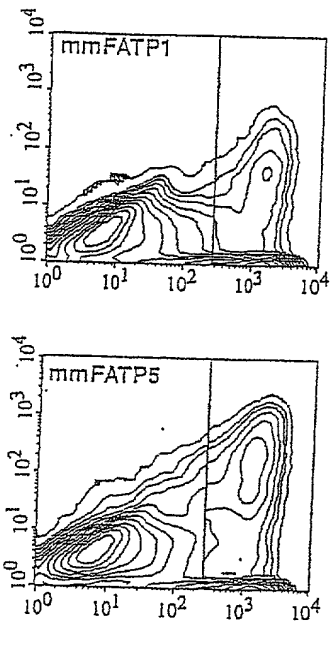
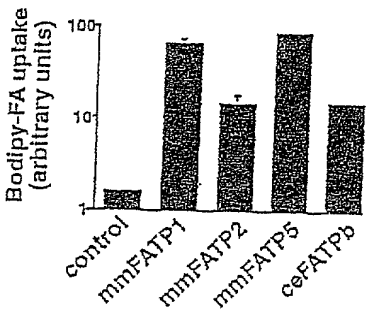


Fig. 2D

Fig. 3



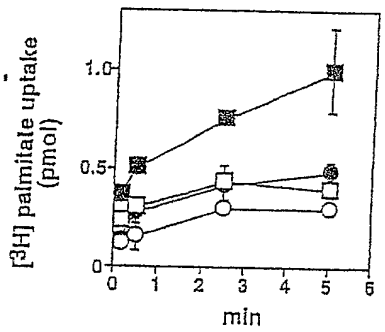


Fig. 4

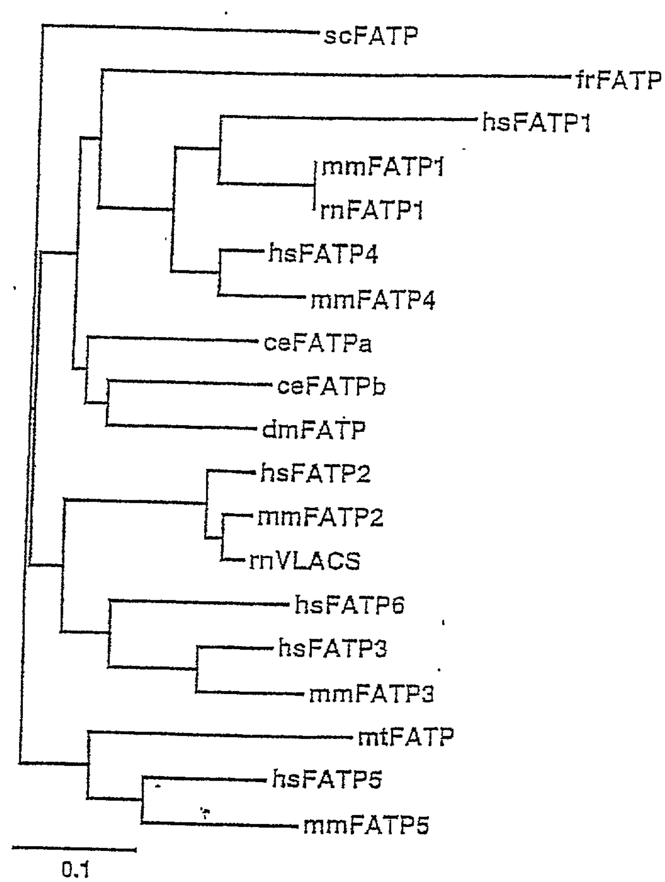
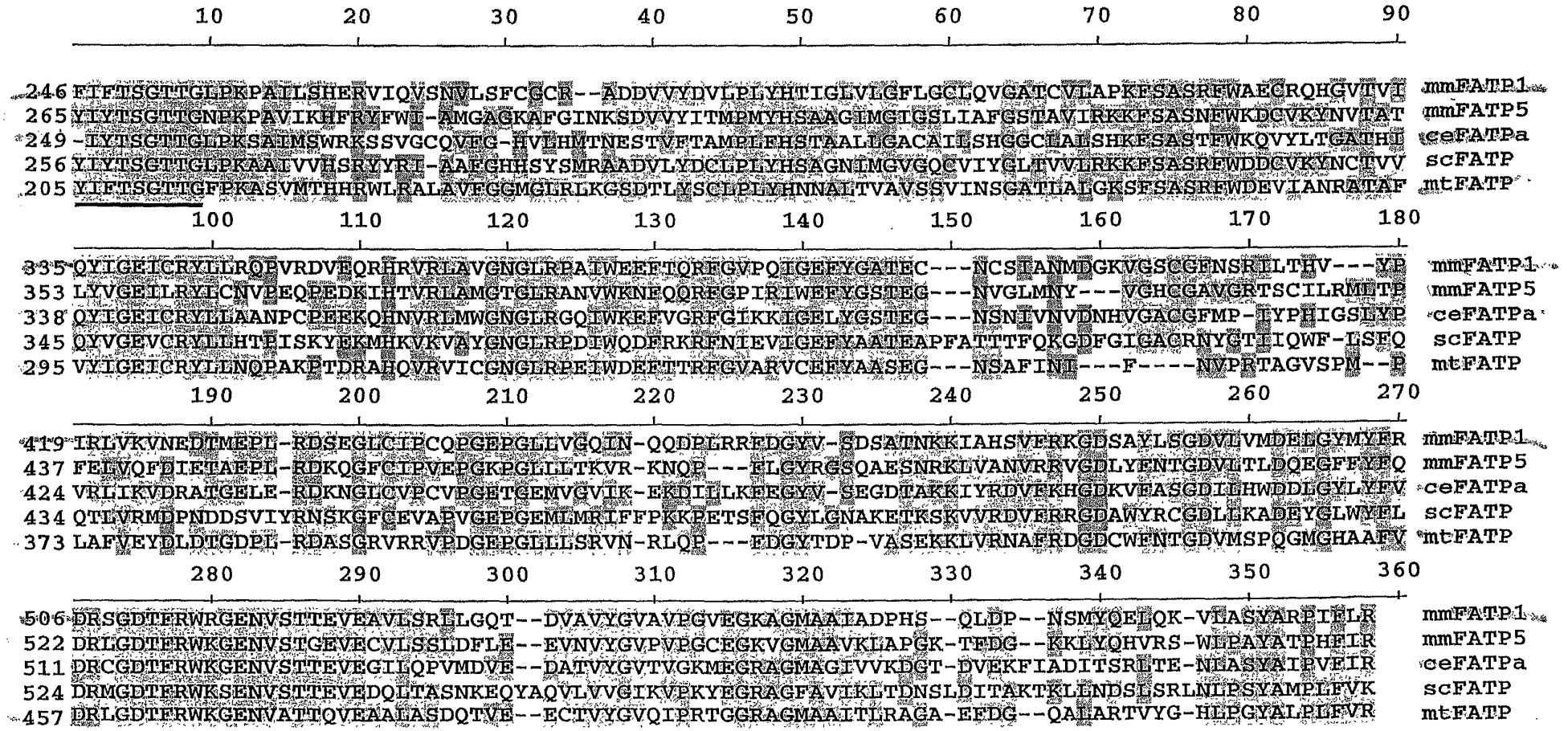

scFATP
frFATP
hsFATP1
mmFATP1
rnFATP1
hsFATP4
mmFATP4
ceFATPa
ceFATPb
dmFATP
hsFATP2
mmFATP2
rnVLACS
hsFATP6
hsFATP3
mmFATP3
mtFATP
hsFATP5
mmFATP5
0.1

Figure 5

FIG. 6



	mmFATP1	mmFATP2	mmFATP3	mmFATP4	mmFATP5	rnFATP1	rnVLACS	ceFATPa	ceFATPb	scFATP	mtFATP
mmFATP1	100.0										
mmFATP2	50.1	100.0									
mmFATP3	51.8	53.3	100.0								
mmFATP4	72.6	48.4	51.7	100.0							
mmFATP5	50.2	54.0	54.4	50.8	100.0						
rnFATP1	97.5	49.7	51.6	71.4	50.4	100.0					
rnVLACS	50.7	95.2	53.0	50.9	53.6	50.3	100.0				
ceFATPa	54.1	48.6	47.2	53.6	47.0	53.9	49.1	100.0			
ceFATPb	55.7	48.8	46.8	54.2	47.3	54.9	49.0	60.3	100.0		
scFATP	45.5	43.2	42.3	47.0	42.7	45.5	43.9	45.3	43.1	100.0	
mtFATP	47.5	47.8	43.9	48.2	47.2	47.1	48.3	44.9	44.5	42.3	100.0

Figure 7

mmFATP3 DNA sequence

```
ACGACTCACTATAGGGAGAGAGCTAATGACGTCGCATGCAC 40
GCGTAAGCTTGGGCCCCTCGAGGGATCCTCTAGAGCGGCC 80
GCCGACCCCGAAAGCTCTGAGAGCGGGTGCAGTCTGGCCT 120
GGCGTCTCGCGTACCTGGCCCCGGAGCAGCCGACACACAC 160
CTTCCTCATCCACGGCGCGCAGCGCTTTAGCTACGCGGAG 200
GCTGAGCGCGAGAGCAACCGGATTGCTCGCGCCTTTCTGC 240
GCGCACGGGGCTGGACCGGGGCCCGCCGAGGCTGGGCAG 280
GGGCAGCACTGAGGAAGGCGCACGCGTGGCGCCTCCGGCT 320
GGAGATGCGGCTGCTAGAGGACGACCGCGCCCCCTCTGG 360
CACCCGGGGCGACCGTGGCGCTGCTCCTCCCAGCGGGCCC 400
```

Figure 8A

GCATTTCCTTTGCATTTGGTTCGGACTGGCCAAAGCTGGC 440
CTGCGCACGGCCTTTGTGCCCACCGCTTTACGCCGAGGAC 480
CCCTGCTGCACTGCCTCCGCAGCTGCGGTGCCGAGTGCGCT 520
CGTGCTGGCCACAGAGTTCCTGGAGTCCCTGGAGCCGGAC 560
CTGCCGGCCTTGAGAGCCATGGGGCTCCACCTATGGCCGA 600
CGGGCCCTGAAACTAATGTAGCTGGAATCAGCAATTTGCT 640
ATCGGAAGCAGCAGACCAAGTGGATGAGCCAGTGCCGGGG 680
TACCTCTCTGCCCCCCAGAACATAATGGACACCTGCCTGT 720
ACATCTTCACCTCTGGCACTACTGGCCTGCCCAAGGCTGC 760
TCGAATCAGTCATCTGAAGGTTCTACAGTGCCAGGGATTC 800
TACCATCTGTGTGGAGTCCACCAGGAGGACGTGATCTACC 840
TCGCACTCCCACTGTACCACATGTCTGGCTCCCTTCTGGG 880
CATTGTGGGCTGCTTGGGCATTGGGGCCACCGTGGTGCTG 920
AAACCCAAGTTCTCAGCTAGCCAGTTCTGGGACGATTGCC 960
AGAAACACAGGGTGACAGTGTTCCAGTACATTGGGGAGTT 1000
GTGCCCGATACCTCGTCAACCAGCCCCCGAGCAAGGCAGAG 1040
TTTGACCATAAGGTGCGCTTGGCAGTGCGGCAGTGGGTTGC 1080
GCCCAGACACCTGGGAGCGTTTCCTGCGGCGATTTGGACC 1120
TCTGCAGATACTGGAGACGTATGGCATGACAGAGGGCAAC 1160
GTAGCTACGTTCAATTACACAGGACGGCAGGGTGCAGTGG 1200
GGCGAGCTTCCTGGCTTTACAAGCACATCTTCCCCTTCTC 1240
CTTGATTCGATACGATGTCATGACAGGGGAGCCTATTCGG 1280
AATGCCCAGGGGCACTGCATGACCACATCTCCAGGTGAGC 1320
CAGGCCTACTGGTGGCCCCAGTGAGCCAGCAGTCCCCCTT 1360
CCTGGGCTATGCTGGGGCTCCGGAGCTGGCCAAGGACAAG 1400
CTGCTGAAGGATGTCTTCTGGTCTGGGGACGTTTTCTTCA 1440
ATACTGGGACCTCTTGGTCTGTGATGAGCAAGGCTTTTCT 1480
TCACTTCCACGATCGTACTGGAGACACCATCAGGTGGAAG 1520
GGAGAGAATGTGGCCACAACTGAAGTGGCTGAGGTCTTGG 1560
AGACCCTGGACTTCCTTCAGGAGGTGAACATCTATGGAGT 1600
CACGGTGCCAGGGCACGAAGGCAGGGCAGGCATGGCGGCC 1640
TTGGCTCTGCGGCCCCCGCAGGCTCTGAACCTGGTGCAGC 1680
TCTACAGCCATGTTTCTGAGAACTTGCCACCGTATGCCCG 1720
ACCTCGGTTTCTCAGGCTCCAGGAATCTTTGGCCACTACT 1760
GAGACCTTCAAACAGCAGAAGGTTAGGATGGCCAATGAGG 1800
GCTTTGACCCCAGTGTACTGTCTGACCCACTCTATGTTCT 1840
GGACCAAGATATAGGGGCCTACCTGCCCCTCACACCTGCC 1880
CGGTACAGTGCCCTCCTGTCTGGAGACCTTCGAATCTGAA 1920
ACCTTCCACTTGAGGGAGGGGCTCGGAGGGTACAGGCCAC 1960
CATGGCTGCACCAGGGAGCGGTTTTCGGGTATCTTTTGTAT 2000
ATGGAGTCATTATTTTGTAATAAACAGCTGGAGCTTAAAA 2040
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2080
AAAAAAA 2087

Figure 8B

mmFATP3 protein sequence

```
AADPESSESGCSLAWRLAYLAREQPIHIFLLHGAQRFSYAEAERESNRIA 50
RAFLRARGWIGGRRGSGRGSIEEGARVAPPAGDAAARGITAPPLAPGATV 100
ALLLPAGPDFLWIWFGLAKAGLRIAFVPTALRRGPLLHCLRSCGASALVL 150
ATEFLESLEPDLPALRAMGLHLWATGPEINVAGISNLLSEAADQVDEPVP 200
GYLSAPQNIMDICLYIFISGITGLPKAARISHLKVLQCQGFYHLCGVHQE 250
DVIYLALPLYHMSGSLLGIVGCLGIGATVVLKPKFSASQFWDDCQKHRVT 300
VFQYIGELCRYLVNQPPSKAEFDHKVRLAVGSGLRPDIWERFLRRFGPLQ 350
ILETYGMTEGNVATFNYTGRQGAVGRASWLYKHIFPFSLIRYDVMIGEPI 400
RNAQGHCMTTSPGEPGLLVAPVSQQSPFLGYAGAPELAKDKLLKDVFWSG 450
DVFFNTGDLLVCDEQGFLHFHDRTGDTIRWKGENVATTEVAEVLEILDFL 500
QEVNIYGVTVPGHEGRAGMAALALRPPQALNLVQLYSHVSENLPPYARPR 550
FLRLQESLATTEIFKQQKVRMANEGFDPSVLSDPLYVLDQDIGAYLPLTP 600
ARYSALLSGDLRI 613
```

Figure 9

mmFATP4 DNA sequence

```
CCCACGCGGTCCGCCCACGCGTCCGGCATGGCCAAGCTGGG 40
CGTGGAGGCGGCTCTCATCAACACCAACCTTAGGCGGGAT 80
GCCCTGCGCCACTGTCTTGACACCTCAAAGGCACGAGCTC 120
TCATCTTTGGCAGTGAGATGGCCTCAGCTATCTGTGAGAT 160
CCATGCTAGCCTGGAGCCCACACTCAGCCTCTTCTGCTCT 200
GGATCCTGGGAGCCCAGCACAGTGCCCGTCAGCACAGAGC 240
ATCTGGACCCTCTTCTGGAAGATGCCCCGAAGCACCTGCC 280
CAGTCACCCAGACAAGGGTTTTACAGATAAGCTCTTCTAC 320
ATCTACACATCGGGCACCACGGGGCTACCCAAAGCTGCCA 360
TTGTGGTGCACAGCAGGTATTATCGTATGGCTTCCCTGGT 400
GTACTATGGATTCCGCATGCGGCCTGATGACATTGTCTAT 440
GACTGCCTCCCCCCTCTACCACTCAAGCAGGAAACATCGTG 480
GGGATTGGCAGTGCTTACTCCACGGCATGACTGTGGTGAT 520
CCGGAAGAAGTTCTCAGCCTCCCCGTTCTGGGATGATTGT 560
ATCAAGTACAACTGCACAGTGGTACAGTACATTGGCGAGC 600
TCTGCCGCTACCTCCTGAACCAGCCACCCCGTGAGGCTGA 640
GTCTCGGCACAAGGTGCGCATGGCACTGGGCAACCGTCTC 680
CGGCAGTCCATCTGGACGGACTTCTCCAGCCGTTTCCACA 720
```

Figure 10A

TCCCCCAGGTGGCTGAGTTCTATGGGGCCACTGAATGCAA 760
CTGTAGCCTGGGCAACTTTGACAGCCGGGTGGGGGCCTGT 800
GGCTTCAATAGCCGCATCCTGTCCTTTGTGTACCCTATCC 840
GTTTGGTACGTGTCAATGAGGATACCATGGAACTGATCCG 880
GGGACCCGATGGAGTCTGCATTCCCTGTCAACCAGGTCAG 920
CCAGGCCAGCTGGTGGGTCGCATCATCCAGCAGGACCCTC 960
TGCGCCGTTTCGACGGGTACCTCAACCAGGGTGCCAACAA 1000
CAAGAAGATTGCTAATGATGTCTTCAAGAAGGGGGACCAA 1040
GCCTACCTCACTGGTGACGTCCTGGTGATGGATGAGCTGG 1080
GTTACCTGTACTTCCGAGATCGCACTGGGGACACGTTCCG 1120
CTGGAAAGGGGAGAATGTATCTACCACTGAGGTGGAGGGC 1160
ACACTCAGCCGCCTGCTTCATATGGCAGATGTGGCAGTTT 1200
ATGGTGTTGAGGTGCCAGGAACTGAAGCCCGAGCAGGAAT 1240
GGCTGCCGTTGCAAGTCCCATCAGCAACTGTGACCTGGAG 1280
AGCTTTGCACAGACCTTGAAAAAGGAGCTGCCTCTGTATG 1320
CCCGCCCCATCTTCCTGCGCTTCTTGCCTGAGCTGCACAA 1360
GACAGGGACCTTCAAGTTCCAGAAGACAGAGTTGCGGAAG 1400
GAGGGCTTTGACCCATCTGTTGTGAAAGACCCGCTGTTCT 1440
ATCTGGATGCTCGGAAGGGCTGCTACGTTGCACTGGACCA 1480
GGAGGCCTATACCCGCATCCAGGCAGGCGAGGAGAAGCTG 1520
TGATTTCCCCCTACATCCCTCTGAGGGCCAGAAGATGCTG 1560
GATTCAGAGCCCTAGCGTCCACCCCAGAGGGTCCTGGGCA 1600
ATGCCAGACCAAAGCTAGCAGGGCCCGCACCTCCGCCCCT 1640
AGGTGCTGATCTCCCCTCTCCCAAACTGCCAAGTGACTCA 1680
CTGCCGCTTCCCCGACCCTCCAGAGGCTTTCTGTGAAAGT 1720
CTCATCCAAGCTGTGTCTTCTGGTCCAGGCGTGGCCCCTG 1760
GCCCCAGGGTTTCTGATAGGCTCCTTTAGGATGGTATCTT 1800
GGGTCCAGCGGGCCAGGGTGTGGGAGAGGAGTCACTAAGA 1840
TCCCTCCAATCAGAAGGGAGCTTACAAAGGAACCAAGGCA 1880
AAGCCTGTAGACTCAGGAAGCTAAGTGGCCAGAGACTATA 1920
GTGGCCAGTCATCCCATGTCCACAGAGGATCTTGGTCCAG 1960
AGCTGCCAAAGTGTCACCTCTCCCTGCCTGCACCTCTGGG 2000
GAAAAGAGGACAGCATGTGGCCACTGGGCACCTGTCTCAA 2040
GAAGTCAGGATCACACACTCAGTCCTTGTTTCTCCAGGTT 2080
CCCTTGTTCTTGTCTCGGGGAGGGAGGGAGGAGTGTCCTG 2120
TCTGTCCTTCCTGCCTGTCTGTGAGTCTGTGTTGCTTCTC 2160
CATCTGTCCTAGCCTGAGTGTGGGTGGAACAGGCATGAGG 2200
AGAGTGTGGCTCAGGGGCCAATAAACTCTGCCTTGACTCC 2240
TCTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2280
AAAAAAAAAAAAAAAAAAAA 2301

Figure 10B

mmFATP4 protein sequence

```
HASAHASGMAKLGVEAALINLNLRRDALRHCLDISKARAL 40
IFGSEMASAICEIHASLEPILSLFCSGSWEPSIVPVSIEH 80
LDPLLEDAPKHLPSHPDKGFIDKLFYIYTSGTTGLPKAAI 120
VVHSRYYRMASLVYYGFRMRPDDIVYDCLPLYHSSRKHRG 160
IWQCLLHGMIVVIRKKFSASRFWDDCIKYNCIVVQYIGEL 200
CRYLLNQPPREAESRHKVRMALGNGLRQSIWIDFSSRFHI 240
PQVAEFYGATECNCSLGNFDSRVGACGFNSRILSFVYPIR 280
LVRVNEDIMELIRGPDGVCIPCQPGQPGQLVGRIIQQDPL 320
RRFDGYLNQGANNKKIANDVFKKGDQAYLTGDVLVMDELG 360
YLYFRDRIGDTFRWKGENVSTFEVEGILSRLLHMADVAVY 400
GVEVPGTEGRAGMAAVASPISNCDLESFAQTLKKELPLYA 440
RPIFLRFLPELHKTGTFKFQKTELRKEGFDPSVVKDPLFY 480
LDARKGCYVALDQEAYTRIQAGEEKL 507
```

Figure 11

mmFATP5 DNA sequence

```
CACTCATCAGAGCTAAGAGAGACTACACGCTCTCATCTAC 40
TTCAGAAAGAGCCAATGCCATGGGTATTTGGAAGAAACTA 80
ACCTTACTGCTGTTGCTGCTTCTGCTGGTTGGCCTGGGGC 120
AGCCCCCATGGCCAGCAGCTATGGCTCTGGCCCTGCGTTG 160
GTTCCTGGGAGACCCCACATGCCTTGTGCTGCTTGGCTTG 200
GCATTGCTGGGCAGACCGTGGATCAGCTCCTGGATGCCCC 240
ACTGGCTGAGCCTGGTAGGAGCAGCTCTTACCTTATTCCT 280
ATTGCCTCTACAGCCACCCCCAGGGCTACGCTGGCTGCAT 320
AAAGATGTGGCTTTCACCTTCAAGATGCTTTTCTATGGCC 360
TAAAGTTCAGGCGACGCCTTAACAAACATCCTCCAGAGAC 400
CTTTGTGGATGCTTTAGAGCGGCAAGCACTGGCATGGCCT 440
GACCGGGTGGCCTTGGTGTGTACTGGGTCTGAGGGCTCCT 480
CAATCACAAATAGCCAGCTGGATGCCAGGTCCTGTCAGGC 520
AGCATGGGTCCTGAAAGCAAAGCTGAAGGATGCCGTAATC 560
CAGAACACAAGAGATGCTGCTGCTATCTTAGTTCTCCCGT 600
CCAAGACCATTTCTGCTTTGAGTGTGTTTCTGGGGTTGGC 640
CAAGTTGGGCTGCCCTGTGGCCTGGATCAATCCACACAGC 680
CGAGGGATGCCCTTGCTACACTCTGTACGGAGCTCTGGGG 720
CCAGTGTGCTGATTGTGGATCCAGACCTCCAGGAGAACCT 760
GGAAGAAGTCCTTCCCAAGCTGCTAGCTGAGAACATTCAC 800
```

Figure 12A

TGCTTCTACCTTGGCCACAGCTCACCCACCCCCGGGAGTAG 840
AGGCTCTGGGAGCTTCCCTGGATGCTGCACCTTCTGACCC 880
AGTACCTGCCAGCCTTCGAGCTACGATTAAGTGGAAATCT 920
CCTGCCATATTCATCTTTACTTCAGGGACCACTGGACTCC 960
CAAAGCCAGCCATCTTATCACATGAGCGGGTCATACAAGT 1000
GAGCAACGTGCTGTCCTTCTGTGGATGCAGAGCTGATGAT 1040
GTGGTCTATGACGTCCTACCTCTGTACCATACGATAGGGC 1080
TTGTCCTTGGATTCCTTGGCTGCTTACAAGTTGGAGCCAC 1120
CTGTGTCCTGGCCCCCAAGTTCTCTGCCTCCCGATTCTGG 1160
GCTGAGTGCCGGCAGCATGGCGTAACAGTGATCTTGTATG 1200
TGGGTGAAATCCTGCGGTACTTGTGTAACGTCCCTGAGCA 1240
ACCAGAAGACAAGATACATACAGTGCGCTTGGCCATGGGA 1280
ACTGGACTTCGGGCAAATGTGTGCGAAAAACTTCCAGCAAC 1320
GCTTTGGTCCCATTCGGATCTGGCAATTCTACGGATCCAC 1360
AGAGGGCAATGTGGGCTTAATGAACTATGTGGGCCACTGC 1400
GGGGCTGTGGGAAGGACCAGCTGCATCCTTCGAATGCTGA 1440
CTCCCTTTGAGCTTGTACAGTTCGACATAGAGACAGCAGA 1480
GCCTCTGAGGGACAAACAGGGTTTTTGCATTCCTGTGGAG 1520
CCAGGAAAGCCAGGACTTCTTTTGACCAAGGTTCGAAAGA 1560
ACCAACCCTTCCTGGGCTACCGTGGTTCCCAGGCCGAGTC 1600
CAATCGGAAACTTGTTGCGAATGTACGACGCGTAGGAGAC 1640
CTGTACTTCAACACTGGGGACGTGCTGACCTTGGACCAGG 1680
AAGGCTTCTTCTACTTTCAAGACCGCCTTGGTGACACCTT 1720
CCGGTGGAAGGGCGAAAACGTATCTACTGGAGAGGTGGAG 1760
TGTGTTTGTCTAGCCTAGACTTCCTAGAGGAAGTCAATG 1800
TCTATGGTGTGCCTGTGCCAGGGTGTGAGGGTAAGGTTGG 1840
CATGGCTGCTGTGAAACTGGCTCCTGGGAAGACTTTTGAT 1880
GGGCAGAAGCTATACCAGCATGTCGCTCCTGGCTCCCTG 1920
CCTATGCCACACCTCATTTCATCGGTATCCAGGATTCCCT 1960
GGAGATCACAAACACCTACAAGCTGGTAAAGTCACGGCTG 2000
GTGCGTGAGGGTTTTGATGTGGGGATCATTGCTGACCCCC 2040
TCTACATACTGGACAACAAGGCCCAGACCTTCGGAGTCT 2080
GATGCCAGATGTGTACCAGGCTGTGTGTGAAGGAACCTGG 2120
AATCTCTGACCACCTAGCCAACTGGAAGGCAATCCAAAAG 2160
TGTAGAGATTGACACTAGTCAGCTTCACAAAGTTGTCCGG 2200
GTTCGAGATGCCCATGGCCCAGTAGTACTTAGAGAATAAA 2240
CTTGAATGTGTATACAAAAAAAAAAAAAAAAAAAAA 2277

Figure 12B

mmFATP5 protein sequence

```
MALALRWFLGDPTCLVLLGLALLGRPWISSWMPHWLSLVG 40
AALLTLFLLPLQPPPGLRWLHKDVAFTFKMLFYGLKFRRRL 80
NKHPPETFVDALERQALAWPDRVALVCTGSEGSSTTNSQL 120
DARSCQAAWVLKAKLKDAVIQNTRDAAAILVLPSKTISAL 160
SVFLGLAKLGCPVAWINPHSRGMPLLHSVRSSGASVLIVD 200
PDLQENLEEVLPKLLAENIHCFYLGHSSPTPGVEALGASL 240
DAAPSDPVPASLRATIKWKSPAIFIFTSGTTGLPKPAILS 280
HERVIQVSNVLSFCGCRADDVVYDVLPLYHTTGLVLGFLG 320
CLQVGATCVLAPKFSASRFWAECRQHGVIVILYVGEILRY 360
LCNVPEQPEDKIHIVRLAMGTGLRANVWKNFQQRFGPIRI 400
WEFYGSTEGNVGLMNYVGHCGAVGRTSCILRMLTPFELVQ 440
FDIETAEPLRDKQGFCIPVEPGKPGLLLTKVRKNQPFLGY 480
RGSQAESNRKLVANVRRVGDLYFNTGDVLTLDQEGFFYFQ 520
DRLGDTFRWKGENVSTGEVECVLSSLDFLEEVNVYGVPVP 560
GCEGKVGMAAVKLAPGKTFDGQKLYQHVRSWLPAYATPHF 600
IRIQDSLETTNTYKLVKSRLVREGFDVGIIADPLYILDNK 640
AQTFRSLMPDVYQAVCEGIWNL 663
```

Figure 13

hsFATP2 DNA sequence

```
ATGGGATTGACTCTTTTCCTGGACAAAGTGGATGAAGTATC 40
AACTGAACCTATCCCAGAGTCATGGAGGTCTGAAGTCACT 80
TTTTCCACTCCTGCCTTATACATTTATACTTCTGGAACCA 120
CAGGTCTTCCAAAAGCAGCCATGATCACTCATCAGCGCAT 160
ATGGTATGGAACTGGCCTCACTTTTGTAAGCGGATTGAAG 200
GCAGATGATGTCATCTATATCACTCTGCCCTTTTACCACA 240
GTGCTGCACTACTGATTGGCATTCACGGATGTATTGTGGC 280
TGGTGCTACTCTTGCCTTGCGGACTAAATTTTCAGCCAGC 320
CAGTTTTGGGATGACTGCAGAAAATACAACGTCACTGTCA 360
TTCAGTATATCGGTGAACTGCTTCGGTATTTATGCAACTC 400
ACCACAGAAACCAAATGACCGTGATCATAAAGTGAGACTG 440
GCACTGGGAAATGGCTTACGAGGAGATGTGTGGAGACAAT 480
TTGTCAAGAGATTTGGGGACATATGCATCTATGAGTTCTA 520
TGCTGCCACTGAAGGCAATATTGGATTTATGAATTATGCG 560
AGAAAAGTTGGTGCTGTTGGAAGAGTAAACTACCTACAGA 600
AAAAAATCATAACTTATGACCTGATTAAATATGATGTGGA 640
GAAAGATGAACCTGTCCGTGATGAAAATGGATATTGCGTC 680
AGAGTTCCCAAAGGTGAAGTTGGACTTCTGGTTTGCAAAA 720
TCACACAACTTACACCATTTAATGGCTATGCTGGAGCAAA 760
GGCTCAGACAGAGAAGAAAAAACTGAGAGATGTCTTTAAG 800
```

Figure 14A

AAAGGAGACCICTATTICAACAGIGGAGATCICTTAATGG 840
TIGACCATGAAAATTICATCTATTICCACGACAGAGTTGG 880
AGATACATTCCGGTGGAAAGGGGAAAATGTGGCCACCACT 920
GAAGTTGCTGATATAGTTGGACTGGTTGATTTTTTTCCAA 960
GGAAGTAAAATGTTTATGGGAGTGCATGGGCCAAGATNAT 1000
GGAGGTTCGAATTGGCATGGCNTTCCNTTCAAAATGGAAA 1040
GAAAACCATGGAATTTGATGGAAAGAAATTTTTTCAGNAC 1080
ATTGCTGATAACCNACCTAGTTATGCAAGGCCCCGGTTTT 1120
NTAAGAANACAGGACACCATTGAGATCACTGGAATTTTTA 1160
AACAGCGCAAAATGACCTTTGGTGGAGGAGGGCTTTAACC 1200
CNGCTGTCATCAAAGATGCCTTGTATTTTCTTGGATGACA 1240
CAGCAAAAATGTATGTGCCTATGACTGAGGACATNTATAA 1280
TGCCATAAGTGNTAAAACCCTGAAATTNTGAATATTCCCA 1320
GGAGGATAATTCAACATTTCCAGAAAGAAACTGAATGGAC 1360
AGCCACTTGATATAATCCAACTTTAATTTGATTGAAGATT 1400
GTCAGGAAATTTTGTAGGAAATTTGCATACCCGTAAAGGG 1440
AGACTTTTTTAAATAACAGTTGAGTCTTTGCAAGTAAAAA 1480
GATTTAGAGATTATTATTTTTCAGTGTGCACCTACTGTTT 1520
GTATTTGCAAACTGAGCTTGTTGGAGGCAAGGCATTATTT 1560
TTTAAAATACTTAGTAAATTAAAGAACACCAACATGTGAA 1600
AAAAAAAAAAAAAAAAAAAAAA 1622

Figure 14B

hsFATP2 protein sequence

YIYISGITGLPKAAMITHQRIWYGIGLITFVSGLKADDVIY 40
ITLPFYHSAALLIGIHGCIVAGATLALRIKFSASQFWDDC 80
RKYNVIVIQYIGELLRYLCNSPQKPNDRDHKVRLALGNGL 120
RGDVWRQFVKRFGDICIYEFYAATEGNIGFMNYARKVGAV 160
GRVNYLQKKLITYDLIKVDVEKDEPVRDENGYCVRVPKGE 200
VGLLVCKITQLTPFNGYAGAKAQIEKKKLRDVFKKGDLYF 240
NSGDLLMVDHENFIYFHDRVGDTFRWKGENVATTEVADIV 280
GLVDFF 286

Figure 15

hsFATP3 DNA sequence

CAATTCGGGACCCCCAGGGGCACTGTATGGCCACATCTCC 40
AGGTGAGCCAGGGGAAGTTGCTAAAGGATGTCTTCCGCCC 80
TGGGGATGTTTTCTTCAACACTGGGGACCTGCTGGTCTGC 120
GATGACCAAGGTTTTCTCCGCTTCCATGATCGTACTGGAG 160

Figure 16A

ACACCTTCAGGTGGAAAGGGCAGAATGTGGCCACAACCGA 200
GGTGGCAGAGGTCTTCGAGGCCCTAGATTTTCTTCAGGAG 240
GTGAACGTCTATGGAGTCACTGTGCCAGGGCATGAAGGCA 280
GGGCTGGAATGGCAGCCCTAGTTCTGCGTCCCCCCCACGC 320
TTTGGACCTTATGCAGCTCTACACCCACGTGTCTGAGAAC 360
TTGCCACCTTATGCCCGGCCCCGATTCCTCAGGCTCCAGG 400
AGTCTTTGGCCACCACAGAGACCTTCAAACAGCAGAAAGT 440
TCGGATGGCAAATGAGGGCTTCGACCCCAGCACCCTGTCT 480
GACCCACTGTACGTTCTGGACCAGGCTGTAGGTGCCTACC 520
TGCCCCCTCACAACTGCCCGGTACAGCGCCCTCCTGGCAGG 560
AAACCTTCGAATCTGAGAACTTCCACACCTGAGGCACCTG 600
AGAGAGGAACTCTGTGGGGTGGGGGCCGTTGCAGGTGTAC 640
TGGGCTGTCAGGGATCTTTTCTATACCAGAACTGGGGTCA 680
CTATTTTGTAATAAATGTGGCTGGAGCTGATCCAGCTGTC 720
TCTGACCTACAAAAAAAAAAAAAAAAAAAAA 753

Figure 16B

hsFATP3 protein sequence

QFGTPRGIVWPHLQVSQGKLLKDVFRPGDVFFNIGDLLVC 40
DDQGFLRFHDRTGDTFRWKGENVATTEVAEVFEALDFLQE 80
VNVYGVTVPGHEGRAGMAALVLRPPHALDLMQLYTHVSEN 120
LPPYARPRFLRLQESLATTETFKQQKVRMANEGFDPSTLS 160
DPLYVLDQAVGAYLPLTTARYSALLAGNLRI 191

Figure 17

hsFATP4 DNA sequence

TCAAGTACAACTGCACGATTGTCATANCATTGGTGAACTG 40
TGCCGNTACCTCCTGAACCAGCCACCGCGGGAGGCAGAAA 80
ACCAGCACCAGGTTCGCATGGCACTAGGCAATGGCCTCCG 120
GCAGTCCATCTGGACCAACTTTTCCAGCCGCTTCCACATA 160
CCCCAGGTGGCTGAGTTYTACGGGGCCACAGAGTGCAACT 200
GTAGCCTGGGCAACTTCGACAGCCAGGTGGGGGCCTGTGG 240
TTTCAATAGCCGCATCCTGTCCTTCGTGTACCCCATCCGG 280
TTGGTACGTGTCAACGAGGACACCATGGAGCTGATCCGGG 320
GGCCGGACGGCGTCTGCATTCCCTGCCAGCCAGGTGAGCC 360
GGGCCAGCTGGTGGGCCGCATCATCCAGAAAGACCCCCTG 400
CGCCGCTTCGATGGCTACCTCAACCAGGGCGCCAACAACA 440
AGAAGATTGCCAAGGATGTCTTCAAGAAGGGGGACCAGGC 480
CTACCTTACTGGTGATGTGCTGGTGATGGACGAGCTGGGC 520

Figure 18A

TACCTGTACTTCCGACACGCACTGGGGACACGTTCCGCT 560
GCAAAGGTGAGAACGTGTCCACCACCGAGGTGGAAGGCAC 600
ACTCAGCCGCCTGCTGGACATGGCTGACGTGGCCGTGTAT 640
GGTGTCGAGGTGCCAGGAACCGAGGCCGGGCCCGAATGG 680
CTGCTGTGGCCAGCCCCACTGGCAACTGTGACCTGGGAGC 720
GCTTTGCTCAGGTC 734

Figure 18B

hsFATP4 protein sequence

IGELCRYLLNQPPREAENQHQVRMALGNGLRQSIWINFSS 40
RFHIPQVAEFYGAIECNCSLGNFDSQVGACGFNSRILSFV 80
YPIRLVRVNEDIMELIRGPDGVCIPCQPGEPGQLVGRLIQ 120
KDPLRRFDGYLNQGANNKKIAKDVFKKGDQAYLTGDVLVM 160
DELGYLYFRDRIGDTFRWKGENVSTTEVEGTLSRLLDMAD 200
VAVYGVEVPGTEG 213

Figure 19

hsFATP5 DNA sequence

CNTGCCTCTTGTACCACGTGATGGGACTTTGTCGTTGGGA 40
TCCTCGGCTGCTTAGATCTCGGAGCCACCTGTGTTCTGGC 80
CCCCAAGTTCTCTACTTCCTGCTTCTGGGATGACTGTCGG 120
CAGCATGGCGTGACAGTGATCCTGTATGTGGGCGAGCTCC 160
TGCGNTACTTGTGTAACATTCCCCAGCAACCAGAGGACCG 200
GACACATACAGTCCGCCTGGCAATGGGCAATGGACTACGG 240
GCTGATGTGTGGGGAGACCTTCCAGCAGCGTTTCGGTCCT 280
ATTTCGGATCTNGGGAAGTCTTACGGGCTYCCACAGAAGG 320
GCAACATGGGGCTTTAGTTCAACTATTGTTGGGGGCGCTG 360
CGGGGSCCTGGRGGCAAAGATGGAGCTTGCCTCCTCCGAA 400
TGCTGTCCCCCTTTGAGCTGGTGCAGTTCGACATGGAGGC 440
GGCGGAGCCTGTGAGGGACAATCAGGGCTTCTGCATCCCT 480
GTAGGGCTAGGGGAGCCGGGGCTGCTGTTGACCAAGGTGG 520
TAAGCCAGCAACCCTTCGTGGGCTACCGCGGCCCCCGAGA 560
GCTGTCGGAACGGAAGCTGGTGCGCAACGTGCGGCAATCG 600
GGCGACGTTTACTACAACACCGGGGACGTACTGGCCATGG 640
ACCGCGAAGGCTTCCTCTACTTCCGCGACCGACTGGGGA 680
CACCTTCCGATGGAAGGGCGAGAACGTGTCCACGCACGAG 720
GTGGAGGGCGTGTTGTCGCAGGTGGACTTCTTGCAACAGG 760
TTAACGTGTATGGCGTGTGCGTGCCAGGTTGTGAGGGTAA 800
GGTGGGCATGGCTGCTGTGGCATTAGCCCCCGGCCAGACT 840

Figure 20A

TTCGACGGGGAGAAGTTGTACCAGCACGTTCGGGCTTGGC 880
TCCCTGCCTACGCTACCCCCCATTTCATCCGCATCCAGGA 920
CGCCATGGAGGTCACCAGCACGTTCAAACTGATGAAGACC 960
CGGTTGGTGCGTGAGGGCTTCAATGTGGGGATCGTGGTTG 1000
ACCCTCTGTTTGTACTGGACAACCGGGCCCAGTCCTTCCG 1040
GCCCCTGACGGCAGAAATGTACCAGGCTGTGTGTGAGGGA 1080
ACCTGGAGGCTCTGATCACCTGGCCAACCCACTGGGGTAG 1120
GGATCAAAGCCAGCCACCCCCACCCCAACACACTCGGTGT 1160
CCCTTTCATCCTGGGCCTGTGTGAATCCCAGCCTGGCCAT 1200
ACCCTCAACCTCAGTGGGCTGGAAATGACAGTGGGCCCTG 1240
TAGCAGTGGCAGAATAAACTCAGMTGYGTTCACAGAAA 1278

Figure 20B

hsFATP5 protein sequence

EGQHGALVQLLLGALRGPGGKDGACLLRMLSPFELVQFDM 40
EAAEPVRDNQGFCIPVGLGEPGLLLTKVVSQQPFVGYRGP 80
RELSERKLVRNVRQSGDVYYNTGDVLAMDREGFLYFRDRL 120
GDTFRWKGENVSTHEVEGVLSQVDFLQQVNVYGVCVPGCE 160
GKVGMAAVALAPGQTFDGEKLYQHVRAWLPAYATPHFIR 199

Figure 21

hsFATP6 DNA sequence

CGCTTGTGTGTTAAAGAAGAAATTTTCAGCAAGCCAGTTT 40
TGGAGTGACTGCAAGAAGTATGATGTGACTGTGTTTCAGT 80
ATATTGGAGAACTTTGTCGCTACCTTTGCAAACAATCTAA 120
GAGAGAAGCAGAAAAGGATCATAAGGTGCGTTTGGCAATT 160
GGAAATGGCATACGGAGTGATGTATGGAGAGAATTTTTAG 200
ACAGATTTGGAAATATAAAGGTGTGTGAACTTTATGCAGC 240
TACCGAATCAAGCATATCTTTCATGAACTACACTGGGAGA 280
ATTGGAGCAATTGGGAGAACAAATTTGTTTTACAAACTTC 320
TTTCCACTTTTGACTTAATAAAGTATGACTTTCAGAAAGA 360
TGAACCCATGAGAAATGAGCAGGGTTGGGTATTCATGAGA 400
AAAAGGAGACCTGGACTTCTCATTTCTCGAGTGAATGCAA 440
AAAATCCCTTCTTTGGCTATGCTGGGCCTTATAAGCACAC 480
AAAAGACAAATTGCTTTGTGATGTTTTTAAGAAGGGAGAT 520
GTTTACCTTAATACTGGAGACTTAATAGTCCAGGATCAGG 560
ACAATTTCCTTTATTTTTGGGACCGTACTGGAGACACTTT 600
CAGATGGAAAGGAGAAAATGTCGCAACCACTGAGGTTGCT 640
GATGTTATTGGAATGTTGGATTTCATACAGGAAGCAAACG 680
TCTATGGTGTGGCTATATCAGGTTATGAAGGAAGAGCAGG 720

Figure 22A

AATGGCTTCTATTATTTTAAAACCAAATACATCTTTAGAT 760
TTGGAAAAAGTTTATGAACAAGTTGTAACATTTCTACCAG 800
CTTATGCTTGTCCACGATTTTTAAGAATTCAGGAAAAAAT 840
GCAAGCAACAGGAACATTCAAACTATTGAAGCATCAGTTG 880
GTGGAAGATGGATTTAATCCACTGAAAATTTCTGAACCAC 920
TTTACTTCATGGATAACTTGAAAAAGTCTTATGTTCTACT 960
GACCAGGGAACTTTATGATCAAATAATGTTAGGGGAAATA 1000
AAACTTTAAGATTTTTATATCTAGAACTTTCATATGCTTT 1040
CTTAGGAAGAGTGAGAGGGGGGTATATGATTCTTTATGAA 1080
ATGGGGAAAGGGAGCTAACATTAATTATGCATGTACTATA 1120
TTTCCTTAATATGAGAGATAATTTTTTAATTGCATAAGAA 1160
TTTTAATTTCTTTTAATTGATATAAACAGAGTTGATTATT 1200
CTTTTTATCTATTTGGAGATTCAGTGCATAACTAAGTATT 1240
TTCCTTAATACTAAAGATTTTAAATAATAAATAGTGGCTA 1280
GCGGTTTGGACAATCACTAAAAATGTACTTTCTAATAAGT 1320
AAAATTTCTAATTTTGAATAAAAGATTAAATTTTACTGAA 1360
A 1361

Figure 22B

hsFATP6 protein sequence

ACVLKKKFSASQFWSDCKKYDVIVFQYIGELCRYLCKQSKREGEKDHKVR 50
LAIGNGIRSDVWREFLDRFGNIKVCELYAATESSISFMNYTGRIGAIGRT 100
NLFYKLLSIFDLIKYDFQKDEFMRNEQGWFMRKRRPGLLISRVNAKNPF 150
FGYAGPYKHIKDKLLCDVFKKGDVYLNIGDLIVQDQLNFLYFWDRIGDIF 200
RWKGENVATTEVADVIGMLDFIQEANVYGVAISGYEGRAGMASIILKENT 250
SLDLEKVYEQVVIFLPAYACPRFLRIQEKMEATGIFKLLKHQLVEDGFNP 300
LKISEPLYFMDNLKKSYVLLITRELYDQIMLGEIKL 335

Figure 23

mtFATP DNA sequence

TAGTCGATAACGTCAAGGAGGCTCTGCGGGGCCTGCGCACC 40
TTCCTGAGGTTGGTCGACAACCAATTCGACATTTCGCAAA 80
CGAATCGAGGGCTTACGTGTCCGATTACTACGGCGGCGCA 120
CACACAACGGTCAGGCTGATCGACCTGGCAACTCGGATGC 160
CGCGAGTGTTGCGGGACACGCCGGTGATTGTGCGTGGGGC 200
AATGACCGGGCTGCTGGCCCGGCGCGAATTCCAAGGCGTCG 240
ATCGGCACGGTGTTCCAGGACCGGGCCGCTCGCTACGGTG 280
ACCGAGTCTTCCTGAAATTCGGCGATCAGCAGCTGACCTA 320
CCGCGACGCTAACGCCACCGCCAACCGGTACGCCGCGGTG 360

Figure 24A

TTGGCCGCCCGCGGCGTCGGCCCCGGCGACGTCGTTGGCA 400
TCATGTTGCGTAACTCACCCAGCACAGTCTTGGCGATGCT 440
GGCCACGGTCAAGTCGGCGCTATCGCCGGCATGCTCAAC 480
TACCACCAGCGCGGCGAGGTGTTGGCGCACAGCCTGGGTC 520
TGCTCGACGCGAAGGTACTGATCGCAGAGTCCGACTTGGT 560
CAGCGCCGTCGCCGAATGCGGCGCCTCGCGCGGCCGGGTA 600
GCGGCGACGTGCTGACCGTCGAGGACGTGGAGCGATTCG 640
CCACAACGGCGCCCGCCACCAACCCGGCGTCGGCGTCGGC 680
GGTGCAAGCCAAAGACACCGCGTTCTACATCTTCACCTCG 720
GGCACCACCGGATTTCCCAAGGCCAGTGTCATGACGCATC 760
ATCGGTGGCTGCGGCGCTGCCGTCTTCGGAGGGATGGG 800
GCTGCGGCTGAAGGGTTCCGACACGCTCTACAGCTGCCTG 840
CCGCTGTACCACAACAACGCGTTAACGGTCGCGGTGTCGT 880
CGGTGATCAATTCTGGGGCGACCCTGGCGCTGGGTAAGTC 920
GTTTTCGGCGTCGCGGTTCTGGGATGAGGTGATTGCCAAC 960
CGGGCGACGGCGTTCGTCTACATCGGCGAAATCTGCCGTT 1000
ATCTGCTCAACCAGCCGGCCAAGCCGACCGACCGTGCCCA 1040
CCAGGTGCGGGTGATCTGCGGTAACGGGCTGCGGCCGGAG 1080
ATCTGGGATGAGTTCACCACCCGCTTCGGGGTCGCGCGGG 1120
TGTGCGAGTTCTACGCCGCCAGCGAAGGCAACTCGGCCTT 1160
TATCAACATCTTCAACGTGCCCAGGACCGCCGGGGTATCG 1200
CCGATGCCGCTTGCCTTTGTGGAATACGACCTGGACACCG 1240
GCGATCCGCTGCGGGATGCGAGCGGGCGAGTGCGTCGGGT 1280
ACCCGACGGTGAACCCGGCCTGTTGCTTAGCCGGGTCAAC 1320
CGGCTGCAGCCGTTCGACGGCTACACCGACCCGGTTGCCA 1360
GCGAAAAGAAGTTGGTGCGCAACGCTTTTCGAGATGGCGA 1400
CTGTTGGTTCAACACCGGTGACGTGATGAGCCCGCAGGGC 1440
ATGGGCCATGCCGCCTTCGTCGATCGGCTGGGCGACACCT 1480
TCCGCTGGAAGGGCGACAATGTCGCCACCACTCAGGTCGA 1520
AGCGGCACTGGCCTCCGACCAGACCGTCGAGGAGTGCACG 1560
GTCTACGGCGTCCAGATTCCGCGCACCGGCGGGCGCGCCG 1600
GAATGGCCGCGATCACACTGCGCGCTGGCGCCGAATTCGA 1640
CGGCCAGGCGCTGGCCCCGAACGGTTTACGGTCACTTGCCC 1680
GGCTATGCACTTCCGCTCTTTGTTCGGGTAGTGGGGTCGC 1720
TGGCGCACACCACGACGTTCAAGAGTCGCAAGGTGGAGTT 1760
GCGCAACCAGGCCTATGGCGCCGACATCGAGGATCCGCTG 1800
TACGTACTGGCCGGCCCGGACGAAGGATATGTGCCGTACT 1840
ACGCCGAATACCCTGAGGAGGTTTCGCTCGGAAGCCGACC 1880
GCAGGGCTAGCGGATTCCGGGCGCAGTCTCGATACCCGCA 1920
CTGGACGCTCGACGGTAACCAGGCACTATGGATGCGTGCG 1960
TTCAACACCGCCGGCCTCAGCCGGTCGTTCAACACCGCCG 2000
GCGTTAG 2007

Figure 24B

mtFATP protein sequence

```
msdyyggahttvrlidlatmmprvladtpvivrgamtgll 40
arpnskasigtvfqdraarygdrvflkfgdqqltyrdana 80
tanryaavlaargvgpgdvvgimlrnspstvlamlatvkc 120
gaiagmlnyhqrgevlahslglldakvliaesdlvsavae 160
cgasrgrvagdvltvedverfattapatmpasasavqakd 200
tafyiftsgttgfpkasvmthhrwlralavfggmglrlkg 240
sdtlysclplyhmnaltvavssvinsgatlalgksfsasr 280
fwdevianratafvyigeicryllngpakptdrahqvrvi 320
cgnglrpeiwdefttrfgvarvcefyaasegnsafinifn 360
vprtagvspmplafveydldtgdplrdasgrvrrvpdgep 400
glllsrvnrlqpfdgytdpvasekklvrnafrdgdcwfnt 440
gdvmspggmghaafvdrlgdtfrwkgenvattqveaalas 480
dqtveectvygvqiprtggragmaaitlragaefdgqala 520
rtvyghlpgyalplfvrvvgslahtttfksrkvelrngay 560
gadiedplyvlagpdegyvpyyaeypeevslgrrpgg 597
```

Figure 25

hsFATP1

```
1       tcg acc cac ggc gtc cgg gac ccc aaa gca gaa gcc cgc aca gta ggc aca gcg cac cca
61      aga agg gtc cag gag tct gca gaa aca gaa agg tcc ccg gcc tca gcc tcc tag tcc ctg
121     cct gcc tcc tgc ctg agc ttc tgg gag act gaa ggc acg gct tgc agc ttc agg atg cgg
                                                                          M   R
181     gct ccg ggt gcg ggc gcg gcc tcg gtg gtc tcg ctg gcg ctg ttg tgg ctg ctg ggg ctg
        A   P   G   A   G   A   A   S   V   V   S   L   A   L   L   W   L   L   G   L
241     ccg tgg acc tgg agc gcg gca gcg gcg ctc ggc gtg tac gtg ggc agc ggc ggc tgg cgc
        P   W   T   W   S   A   A   A   A   L   G   V   Y   V   G   S   G   G   W   R
301     ttc ctg cgc atc gtc tgc aag acc gcg agg cga gac ctc ttc ggt ctc tct gtg ctg atc
        F   L   R   I   V   C   K   T   A   R   R   D   L   F   G   L   S   V   L   I
361     cgc gtg cgc ctg gag ctg cgg cgg cac cag cgt gcc ggc cac acc atc ccg cgc atc ttt
        R   V   R   L   E   L   R   R   H   Q   R   A   G   H   T   I   P   R   I   F
421     cag gcg gta gtg cag cga cag ccc gag cgc ctg gcg ctg gtg gat gcc ggg acc ggc gag
        Q   A   V   V   Q   R   Q   P   E   R   L   A   L   V   D   A   G   T   G   E
481     tgc tgg acc ttt gcg cag ctg gac gcc tac tcc aat gcg gta gcc aac ctc ttc cgc cag
        C   W   T   F   A   Q   L   D   A   Y   S   N   A   V   A   N   L   F   R   Q
541     ctg ggc ttc gcg ccg ggc gac gtg gtg gcc atc ttc ctg gag ggc cgg ccg gag ttc gtg
        L   G   F   A   P   G   D   V   V   A   I   F   L   E   G   R   P   E   F   V
601     ggg ctg tgg ctg ggc ctg gcc aag gcg ggc atg gag gcc gcg ctg ctc aac gtg aac ctg
        G   L   W   L   G   L   A   K   A   G   M   E   A   A   L   L   N   V   N   L
661     cgg cgc gag ccc ctg gcc ttc tgc ctg ggc acc tcg ggc gct aag gcc ctg atc ttt gga
        R   R   E   P   L   A   F   C   L   G   T   S   G   A   K   A   L   I   F   G
721     gga gaa atg gtg gcg gcg gtg gcc gaa gtg agc ggg cat ctg ggg aaa agt ttg atc aag
        G   E   M   V   A   A   V   A   E   V   S   G   H   L   G   K   S   L   I   K
781     ttc tgc tct gga gac ttg ggg ccc gag ggc atc ttg ccg gac acc cac ctc ctg gac ccg
        F   C   S   G   D   L   G   P   E   G   I   L   P   D   T   H   L   L   D   P
841     ctg ctg aag gag gcc tct act gcc ccc ttg gca cag atc ccc agc aag ggc atg gac gat
        L   L   K   E   A   S   T   A   P   L   A   Q   I   P   S   K   G   M   D   D
901     cgt ctt ttc tac atc tac acg tcg ggg acc acc ggg ctg ccc aag gct gcc att gtc gtg
        R   L   F   Y   I   Y   T   S   G   T   T   G   L   P   K   A   A   I   V   V
961     cac agc agg tac tac cgc atg gca gcc ttc ggc cac cac gcc tac cgc atg cag gcg gct
        H   S   R   Y   Y   R   M   A   A   F   G   H   H   A   Y   R   M   Q   A   A
1021    gac gtg ctc tat gac tgc ctg ccc ctg tac cac tcg gca gga aac atc atc ggc gtg ggg
        D   V   L   Y   D   C   L   P   L   Y   H   S   A   G   N   I   I   G   V   G
1081    cag tgt ctc atc tat ggg ctg aca gtc gtc ctc cgc aag aaa ttc tcg gcc agc cgc ttc
        Q   C   L   I   Y   G   L   T   V   V   L   R   K   K   F   S   A   S   R   F
1141    tgg gac gac tgc atc aag tac aac tgc acg gtg gtt cag tac atc ggg gag atc tgc cgc
        W   D   D   C   I   K   Y   N   C   T   V   V   Q   Y   I   G   E   I   C   R
1201    tac ctg ctg aag cag ccg gtg cgc gag gcg gag agg cga cac cgc gtg cgc ctg gcg gtg
        Y   L   L   K   Q   P   V   R   E   A   E   R   R   H   R   V   R   L   A   V
1261    ggg aac ggg ctg cgt cct gcc atc tgg gag gag ttc acg gag cgc ttc ggc gta cgc caa
        G   N   G   L   R   P   A   I   W   E   E   F   T   E   R   F   G   V   R   Q
1321    atc ggg gag ttc tac ggc gcc acc gag tgc aac tgc agc att gcc aac atg gac ggc aag
        I   G   E   F   Y   G   A   T   E   C   N   C   S   I   A   N   M   D   G   K
1381    gtc ggc ttc tgt ggt ttc aac agc cgc atc ctg ccc cac gtg tac ccc atc cgg ctg gtg
        V   G   S   C   G   F   N   S   R   I   L   P   H   V   Y   P   I   R   L   V
1441    aag gtc aat gag gac aca atg gag ctg ctg cgg gat gcc cag ggc ctc tgc atc ccc tgc
        K   V   N   E   D   T   M   E   L   L   R   D   A   Q   G   L   C   I   P   C
1501    cag gcc ggg gag cct ggc ctc ctt gtg ggt cag atc aac caa cag gac ccg ctg cgc cgc
        Q   A   G   E   P   G   L   L   V   G   Q   I   N   Q   Q   D   P   L   R   R
1561    ttc gat ggc tat gtc agc gag agc gcc acc agc aag aag atc gcc cac agc gtc ttc agc
        F   D   G   Y   V   S   E   S   A   T   S   K   K   I   A   H   S   V   F   S
1621    aag ggc gac agc gcc tac ctc tca ggt gac gtg cta gtg atg gat gag ctg ggc tac atg
        K   G   D   S   A   Y   L   S   G   D   V   L   V   M   D   E   L   G   Y   M
1681    tac ttc cgg gac cgt agc ggg gac acc ttc cgc tgg cga ggg gag aac gtc tcc acc acc
        Y   F   R   D   R   S   G   D   T   F   R   W   R   G   E   N   V   S   T   T
1741    gag gtg gag ggc gtg ctg agc cgc ctg ctg ggc cag aca gac gtg gcc gtc tat ggg gtg
        E   V   E   G   V   L   S   R   L   L   G   Q   T   D   V   A   V   Y   G   V
1801    gct gtt cca gga gtg gag ggt aag gca ggg atg gcg gcc gtc gca gac ccc cac agc ctg
        A   V   P   G   V   E   G   K   A   G   M   A   A   V   A   D   P   H   S   L
1861    ctg gac ccc aac gcg ata tac cag gag ctg cag aag gtg ctg gca ccc tat gcc cgg ccc
        L   D   P   N   A   I   Y   Q   E   L   Q   K   V   L   A   P   Y   A   R   P
1921    atc ttc ctg cgc ctc ctg ccc cag gtg gac acc aca ggc acc ttc aag atc cag aag acg
        I   F   L   R   L   L   P   Q   V   D   T   T   G   T   F   K   I   Q   K   T
1981    agg ctg cag cga gag ggc ttt gac cca cgc cag acc tca gac cgg ctc ttc ttc ctg gac
        R   L   Q   R   E   G   F   D   P   R   Q   T   S   D   R   L   F   F   L   D
2041    ctg aag cag ggc cac tac ctg ccc tta aat gag gca gtc tac act cgc atc tgc tcg ggc
        L   K   Q   G   H   Y   L   P   L   N   E   A   V   Y   T   R   I   C   S   G
2101    gcc ttc gcc ctc tga agc tgt tcc tct act ggc cac aaa ctc tgg gcc tgg tgg gag agg
        A   F   A   L   *
2161    cca gct tga gcc aga cag cgc tgc cca ggg gtg gcc gcc tag tac aca ccc acc tgg ccg
2221    agc tgt acc tgg cac ggc cca tcc tgg act gag aaa ctg gaa cct cag agg aac ccg tgc
2281    ctc tct gct gcc ttg gtg ccc ctg tgt ctg cct cct ctc cct gct ttt cag cct ctg tct
2341    cct tcc atc cct gtc cct gtc tgg cct taa ctc ttc cct ctc ttt ctt ttc ttt cct tct
2401    ttc ttt ttt ttt aag ata gag tct cac tct gct gcc cgg gct aga gtg cag tgg tgg gat
2461    ctc ggc tca ctg caa cct ctg cct cct ggg gtt caa gtg atc ctc cca cct cag cct cct
2521    gag tag ctg gga tta cag gca ccc gcc acc acg tcc agc taa ttt tta tat ttt tag tag
2581    aga cgg ggt ttc acc atg ttg gtc agg ctg gtc ttg aac tcc tga cct cag gtg atc cgc
2641    tgg cct cgg cct ccc aga gtg ctg gga tta tag gcg tga gcc tct ggc ccg gcc ttt cct
2701    ttt tcc tct cct ctc ctg ccg aga gtg gaa cac acg tgt cct ggg agc tgc atc ttg tgt
2761    agg gtc cag ctg ctt ttg ggg act gca gga atc atc tcc cct ggg ccc tgg act cgg act
2821    ggg gcc tcc cca cct ccc tct cgg ctg tgc ctt acg gag ccc caa tcc agg cct cct gtg
2881    gct gtt ggg ttc cag atg ctg cag ctc cat gtg act tcc aag cag gcc ctc cgc cct ccc
2941    tgc tga atg gag gag ccg ggg gtc ccc cag gcc aac tgg aaa atc tcc cag gct agg cca
3001    att gcc ttt tgc act tcc ccg ttc ctg tca cat ttc ccc agc ccc acc ttc ccc tcc tga
3061    tgc cct gaa agc ttc cgg aat tga ctg tga cca ctt gga tgt cac cac tgt cag ccc ctg
3121    cct tga tgt ccc cat tta gca atc tcc atg gag ctc ctg ctg gag ggc cct gaa ccc tgc
3181    act gcg tgg ctg ccc agc cag ctg cct cct gtc ctg gga gga ggc ctc ctg ggt gtc ctc
3241    atc tgg tgt gtc tac tgg agg gtc cca cag gag agg cag cag agg ggt cag ggg agg tct
3301    cct gcc ggg ggt tgg cct ctc aag cct cag ggg ttc tag cct gtt gaa tat acc cca cct
3361    ggt ggg tgg ccc ctc cga tgt ccc cac tga tgg ctc tga cac cgt gtt ggt ggc gat gtc
3421    cca gac aat ccc acc agg acg gcc cag aca tcc cta ctg gct tcg ctg gtg gct cat ctc
3481    gaa cat cca cgc cag cct ttc tgg ggc cgg cca ccc agg ccg cct gtc cgt ctg tcc tcc
3541    ctc cag cag cac ccc ctg gcc cct gga gtg gtg ggg cca tgg caa gag aca ccg tgg cgt
3601    ctc atg tga act ttc ctg ggc act gtg gtt tta ttt cct aat tga ttt aag aaa taa acc
3661    tga aga ccg tct ggt gaa aaa aaa aaa aaa aaa agg gcg gcc gc
```

Figure 26

hsFATP4

```
1       cga ccc acg cgt ccg ggc ggg cgg ggc cgg gcg gcg ggc ggg gct ggc ggg gcg gcc ggg
61      cca tgc agg gcg cag agc cgg cta aac cct gct gag acc cgg ctc cgt gcg tcc agg ggc
121     ggc taa tgc ccc tca cgc tgt cta cgc tgc tgc aac cgg gcc gca tct gga cgg ggc gcc
181     gcg cgg cgg agc cga cgc cgg gcc aca atg ctg ctt gga gcc tct ctg gtg ggg gtg ctg
                                        M   L   L   G   A   S   L   V   G   V   L
241     ctg ttc tcc aag ctg gtg ctg aaa ctg ccc tgg acc cag gtg gga ttc tcc ctg ttg ttc
        L   F   S   K   L   V   L   K   L   P   W   T   Q   V   G   F   S   L   L   F
301     ctc tac ttg gga tct ggc ggc tgg cgc ttc atc cgg gtc ttc atc aag acc atc agg cgc
        L   Y   L   G   S   G   G   W   R   F   I   R   V   F   I   K   T   I   R   R
361     gat atc ttt ggc ggc ctg gtc ctc ctg aag gtg aag gca aag gtg cga cag tgc ctg cag
        D   I   F   G   G   L   V   L   L   K   V   K   A   K   V   R   Q   C   L   Q
421     gag cgg cgg aca gtg ccc att ttg ttt gcc tct acc gtt cgg cgc cac ccc gac aag acg
        E   R   R   T   V   P   I   L   F   A   S   T   V   R   R   H   P   D   K   T
481     gcc ctg atc ttc gag ggc aca gat acc cac tgg acc ttc cgc cag ctg gat gag tac tca
        A   L   I   F   E   G   T   D   T   H   W   T   F   R   Q   L   D   E   Y   S
541     agc agt gta gcc aac ttc ctg cag gcc cgg ggc ctg gcc tcg ggc gat gtg gct gcc atc
        S   S   V   A   N   F   L   Q   A   R   G   L   A   S   G   D   V   A   A   I
601     ttc atg gag aac cgc aat gag ttc gtg ggc cta tgg ctg ggc atg gcc aag ctc ggt gtg
        F   M   E   N   R   N   E   F   V   G   L   W   L   G   M   A   K   L   G   V
661     gag gca gcc ctc atc aac acc aac ctg cgg cgg gat gct ctg ctc cac tgc ctc acc acc
        E   A   A   L   I   N   T   N   L   R   R   D   A   L   L   H   C   L   T   T
721     tcg cgc gca cgg gcc ctt gtc ttt ggc agc gaa atg gcc tca gcc atc tgt gag gtc cat
        S   R   A   R   A   L   V   F   G   S   E   M   A   S   A   I   C   E   V   H
781     gcc agc ctg gac ccc tcg ctc agc ctc ttc tgc tct ggc tcc tgg gag ccc ggt gcg gtg
        A   S   L   D   P   S   L   S   L   F   C   S   G   S   W   E   P   G   A   V
841     cct cca agc aca gaa cac ctg gac cct ctg ctg aaa gat gct ccc aag cac ctt ccc agt
        P   P   S   T   E   H   L   D   P   L   L   K   D   A   P   K   H   L   P   S
901     tgc cct gac aag ggc ttc aca gat aaa ctg ttc tac atc tac aca tcc ggc acc aca ggg
        C   P   D   K   G   F   T   D   K   L   F   Y   I   Y   T   S   G   T   T   G
961     ctg ccc aag gcc gcc atc gtg gtg cac agc agg tat tac cgc atg gct gcc ctg gtg tac
        L   P   K   A   A   I   V   V   H   S   R   Y   Y   R   M   A   A   L   V   Y
1021    tat gga ttc cgc atg cgg ccc aac gac atc gtc tat gac tgc ctc ccc ctc tac cac tca
        Y   G   F   R   M   R   P   N   D   I   V   Y   D   C   L   P   L   Y   H   S
1081    gca gga aac atc gtg gga atc ggc cag tgc ctg ctg cat ggc atg acg gtg gtg att cgg
        A   G   N   I   V   G   I   G   Q   C   L   L   H   G   M   T   V   V   I   R
1141    aag aag ttc tca gcc tcc cgg ttc tgg gac gat tgt atc aag tac aac tgc acg att gtg
        K   K   F   S   A   S   R   F   W   D   D   C   I   K   Y   N   C   T   I   V
1201    cag tac att ggt gaa ctg tgc cgc tac ctc ctg aac cag cca ccg cgg gag gca gaa aac
        Q   Y   I   G   E   L   C   R   Y   L   L   N   Q   P   P   R   E   A   E   N
1261    cag cac cag gtt cgc atg gca cta ggc aat ggc ctc cgg cag tcc atc tgg acc aac ttt
        Q   H   Q   V   R   M   A   L   G   N   G   L   R   Q   S   I   W   T   N   F
1321    tcc agc cgc ttc cac ata ccc cag gtg gct gag ttc tac ggg gcc aca gag tgc aac tgt
        S   S   R   F   H   I   P   Q   V   A   E   F   Y   G   A   T   E   C   N   C
1381    agc ctg ggc aac ttc gac agc cag gtg ggg gcc tgt ggt ttc aat agc cgc atc ctg tcc
        S   L   G   N   F   D   S   Q   V   G   A   C   G   F   N   S   R   I   L   S
1441    ttc gtg tac ccc atc cgg ttg gta cgt gtc aac gag gac acc atg gag ctg atc cgg ggg
        F   V   Y   P   I   R   L   V   R   V   N   E   D   T   M   E   L   I   R   G
1501    ccc gac ggc gtc tgc att ccc tgc cag cca ggt gag ccg ggc cag ctg gtg ggc cgc atc
        P   D   G   V   C   I   P   C   Q   P   G   E   P   G   Q   L   V   G   R   I
1561    atc cag aaa gac ccc ctg cgc cgc ttc gat ggc tac ctc aac cag ggc gcc aac aac aag
        I   Q   K   D   P   L   R   R   F   D   G   Y   L   N   Q   G   A   N   N   K
1621    aag att gcc aag gat gtc ttc aag aag ggg gac cag gcc tac ctt act ggt gat gtg ctg
        K   I   A   K   D   V   F   K   K   G   D   Q   A   Y   L   T   G   D   V   L
1681    gtg atg gac gag ctg ggc tac ctg tac ttc cga gac cgc act ggg gac acg ttc cgc tgg
        V   M   D   E   L   G   Y   L   Y   F   R   D   R   T   G   D   T   F   R   W
1741    aaa ggt gag aac gtg tcc acc acc gag gtg gaa ggc aca ctc agc cgc ctg ctg gac atg
        K   G   E   N   V   S   T   T   E   V   E   G   T   L   S   R   L   L   D   M
1801    gct gac gtg gcc gtg tat ggt gtc gag gtg cca gga acc gag ggc cgg gcc gga atg gct
        A   D   V   A   V   Y   G   V   E   V   P   G   T   E   G   R   A   G   M   A
1861    gct gtg gcc agc ccc act ggc aac tgt gac ctg gag cgc ttt gct cag gtc ttg gag aag
        A   V   A   S   P   T   G   N   C   D   L   E   R   F   A   Q   V   L   E   K
1921    gaa ctg ccc ctg tat gcg cgc ccc atc ttc ctg cgc ctc ctg cct gag ctg cac aaa aca
        E   L   P   L   Y   A   R   P   I   F   L   R   L   L   P   E   L   H   K   T
1981    gga acc tac aag ttc cag aag aca gag cta cgg aag gag ggc ttt gac ccg gct att gtg
        G   T   Y   K   F   Q   K   T   E   L   R   K   E   G   F   D   P   A   I   V
2041    aaa gac ccg ctg ttc tat cta gat gcc cag aag ggc cgc tac gtc ccg ctg gac caa gag
        K   D   P   L   F   Y   L   D   A   Q   K   G   R   Y   V   P   L   D   Q   E
2101    gcc tac agc cgc atc cag gca ggc gag gag aag ctg tga ttc ccc cca tcc ctc tga ggg
        A   Y   S   R   I   Q   A   G   E   E   K   L   *
2161    ccg gcg gat gct gga tcc gga gcc cca ggt tcc gcc cca gag cgg tcc tgg aca agg cca
2221    gac caa agc aag cag ggc ctg gca cct cca tcc tga ggt gct gcc cct cca tcc aaa act
2281    gcc aag tga ctc att gcc ttc cca acc ctt cca gag gct ttc tgt gaa agt ctc atg tcc
2341    aag ttc cgt ctt ctg ggc tgg gca ggc cct ctg gtt ccc agg ctg aga ctg acg ggt ttt
2401    ctc agg atg atg tct tgg gtg agg gta ggg aga gga caa ggg gtc acc gag ccc ttc cca
2461    gag agc agg gag ctt ata aat gga acc aga gca gaa gtc ccc aga ctc agg aag tca aca
2521    gag tgg gca ggg aca gtg gta gca tcc atc tgg tgg cca aag aga atc gta gcc cca gag
2581    ctg ccc aag ttc act ggg ctc cac ccc cac ctc cag gag ggg agg aga gga cct gac atc
2641    tgt agg tgg ccc ctg atg ccc cat cta cag cag gag gtc agg acc acg ccc ctg gcc tct
2701    ccc cac tcc ccc atc ctc ctc cct ggg tgg ctg cct gat tat ccc tca ggc agg gcc tct
2761    cag tcc ttg tgg gtc tgt gtc acc tcc atc tca gtc ttg gcc tgg cta tga ggg gag gag
2821    gaa tgg gag agg ggg ctc agg ggc caa taa act ctg cct tga gtc ctc cta aaa aaa aaa
2881    aaa aaa aaa aaa aaa aaa aaa aaa aaa ggg cgg ccg c
```

Figure 27

Protein sequence 646 a.a. MRAPGAGAASVV ... VYTRICSGAFAL

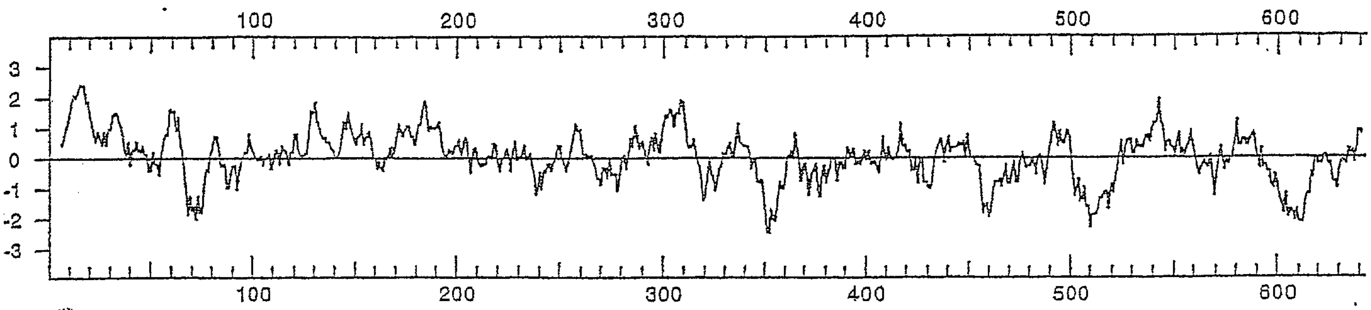


Figure 28A

Protein sequence 646 a.a. MRAPGAGAASVV ... VYTRICSGAFAL

646 Amino Acids MW : 71062 Dalton

		n	n(%)	MW	MW(%)
A ala	alanine	64	9.9	4546	6.4
C cys	cysteine	15	2.3	1545	2.2
D asp	aspartic acid	30	4.6	3450	4.9
E glu	glutamic acid	31	4.8	4000	5.6
F phe	phenylalanine	29	4.5	4264	6.0
G gly	glycine	63	9.8	3592	5.1
H his	histidine	13	2.0	1781	2.5
I ile	isoleucine	29	4.5	3279	4.6
K lys	lysine	22	3.4	2818	4.0
L leu	leucine	77	11.9	8707	12.3
M met	methionine	11	1.7	1441	2.0
N asn	asparagine	15	2.3	1710	2.4
P pro	proline	29	4.5	2814	4.0
Q gln	glutamine	25	3.9	3201	4.5
R arg	arginine	49	7.6	7648	10.8
S ser	serine	33	5.1	2872	4.0
T thr	threonine	27	4.2	2728	3.8
V val	valine	51	7.9	5052	7.1
W trp	tryptophan	9	1.4	1674	2.4
X ukw	unknown	-	-		
Y tyr	tyrosine	24	3.7	3913	5.5
Z ---	STOP	-	-		

Figure 28B

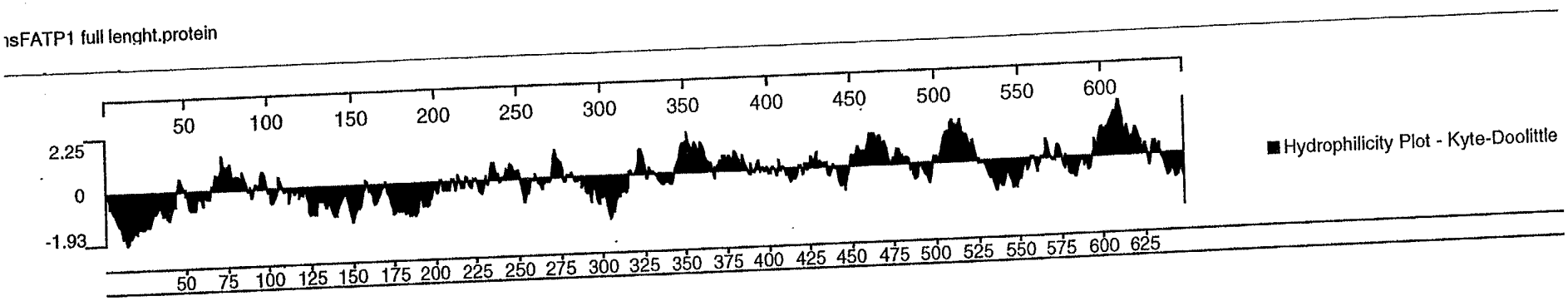


Figure 28C

hsFATP4.pep -> KD Hydrophobicity <11/1>

Protein sequence 643 a.a. MLLGASLVGVLL ... AYSRIQAGEEKL

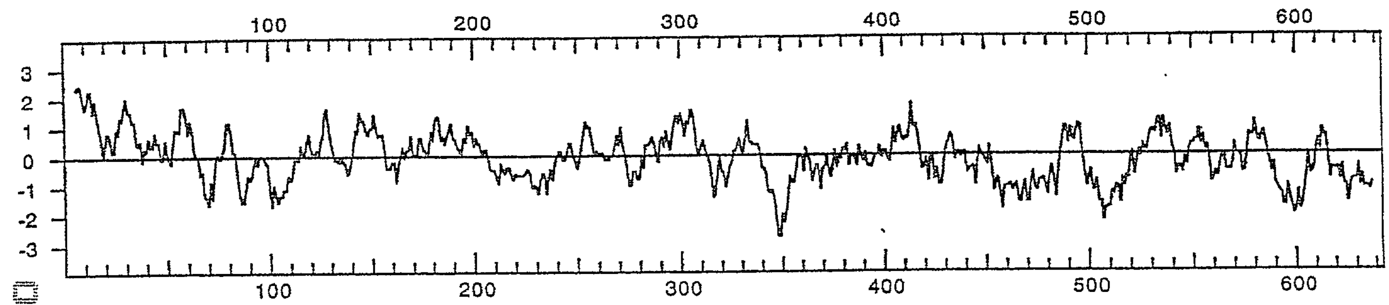


Figure 29A

hsFATP4.pep -> A. A. Usage

Protein sequence 643 a.a. MLLGASLVGVLL ... AYSRIQAGEEKL

643 Amino Acids MW : 72018 Dalton

			n	n(%)	MW	MW(%)
A	ala	alanine	46	7.2	3267	4.5
C	cys	cysteine	16	2.5	1648	2.3
D	asp	aspartic acid	33	5.1	3795	5.3
E	glu	glutamic acid	33	5.1	4258	5.9
F	phe	phenylalanine	34	5.3	5000	6.9
G	gly	glycine	54	8.4	3079	4.3
H	his	histidine	12	1.9	1644	2.3
I	ile	isoleucine	30	4.7	3392	4.7
K	lys	lysine	31	4.8	3970	5.5
L	leu	leucine	76	11.8	8594	11.9
M	met	methionine	12	1.9	1572	2.2
N	asn	asparagine	21	3.3	2394	3.3
P	pro	proline	31	4.8	3008	4.2
Q	gln	glutamine	23	3.6	2945	4.1
R	arg	arginine	45	7.0	7024	9.8
S	ser	serine	35	5.4	3046	4.2
T	thr	threonine	32	5.0	3233	4.5
V	val	valine	46	7.2	4557	6.3
W	trp	tryptophan	8	1.2	1488	2.1
X	ukw	unknown	-	-		
Y	tyr	tyrosine	25	3.9	4076	5.7
Z	---	STOP	-	-		

Figure 29B

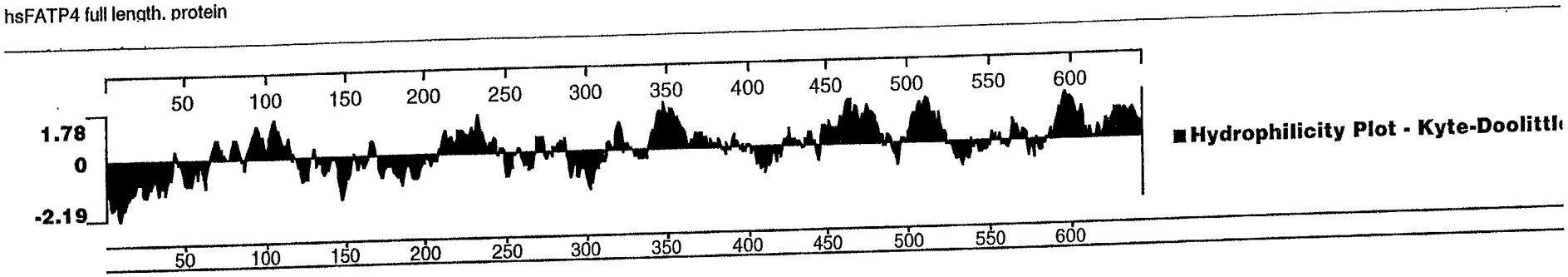


Figure 29C

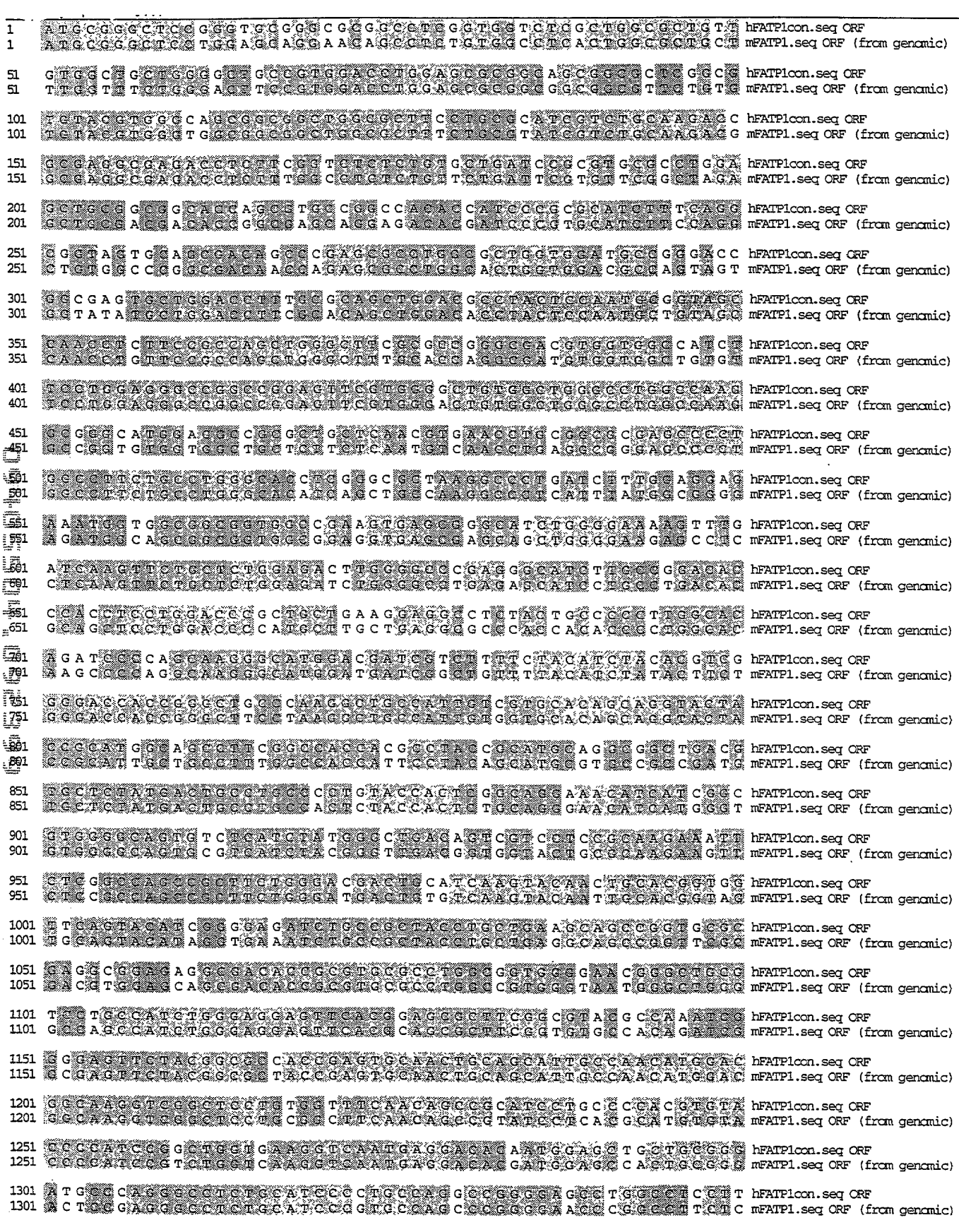

1 ATGCGGGCTCCGGGTGCGGGCGCGGCCTGGGTGCTCTCGCTGGCGCTGTT hFATP1con.seq ORF
1 ATGCGGGCTCCTGGAGCAGGAACAGCCTCTGTGGCCTCACTGGCGCCTGCTT mFATP1.seq ORF (from genomic)
51 GTGGCTGCTGGGGCTGCCGTGGACCTGGAGCGCGGCAGCGGCGCTCGGCG hFATP1con.seq ORF
51 TTGGTTTCTGCGACTCCGTGGACCTGGAGCGCGGCGGCGGCGTCCTGTG mFATP1.seq ORF (from genomic)
101 TCTACGTGGGCAGCGGCGGCTGGCGCTTCCTGCGCATCGTCTGCAAGAGC hFATP1con.seq ORF
101 TCTACGTGGGTGGCGGCGGCTGGCGCTTTCTGCGTATCGTCTGCAAGACG mFATP1.seq ORF (from genomic)
151 GCGAGGCGAGACCTCTTCGGTCTCTCTGTGCTGATCCGCGTGCGCCTGGA hFATP1con.seq ORF
151 GCGAGGCGAGACCTCTTTGCCTCTCTGTTCTGATTCGTGTTCGGCTAGA mFATP1.seq ORF (from genomic)
201 GCTGCGGCGGCACCAGCGTGCCGCCACACCATCCCGCGCATCTTTCAGG hFATP1con.seq ORF
201 GCTGCGACGACACCGCGGAGCAGGAGACACGATCCCGTGCATCTTCCAGG mFATP1.seq ORF (from genomic)
251 CGCTAGTGGAGCGACAGCCCGAGGCGCCTGGCGCTGGTGGATGCCGGGACC hFATP1con.seq ORF
251 CTCTGGCCCGGCGACAACCAGAGCGCCTGGCACTGGTGGACGCCAGTAGT mFATP1.seq ORF (from genomic)
301 GCCGAGTGCTGGACCTTTGCGCAGCTGCAGCGCCTACTGCAATGCGGTAGC hFATP1con.seq ORF
301 GCTATATGCTGGACCTTCGCACAGCTGGAGCACCTACTCCAATGCTGTAGC mFATP1.seq ORF (from genomic)
351 CAACCTCTTCCGCCAGCTGGGCTTCGCGCGGGCGACGTGGTGGCCATCT hFATP1con.seq ORF
351 CAACCTGTTCCGCCAGCTGGGCTTTGCACGAGGCGATGTGGTGGCTGTT mFATP1.seq ORF (from genomic)
401 TCGTGGAGGGCCGGCCGGAGTTCCTGCGGCTGTGGCTGGCCTGGCCAAG hFATP1con.seq ORF
401 TCCTGGAGGGCCGGCCCGAGTTCGTGCGACTGTGGCTGGCCTGGCCAAG mFATP1.seq ORF (from genomic)
451 GCGCGCATGGAGGCCGCGCTGCTCAACGTGAACCTGCGGCGCGAGCCCCT hFATP1con.seq ORF
451 GCCGCTGTGGTGGCTGCTCTCAATGTCAACCTGAGGCGGAGCCCGT mFATP1.seq ORF (from genomic)
501 GCCCTTCTGCCTGGGCACCTCGGCGCTAAGGCCGTGATCTTGCAGGAG hFATP1con.seq ORF
501 GCCCTTCTGCCTGGGCACACCAGCTGCCAAGGCCGTCATTATGCCGGG mFATP1.seq ORF (from genomic)
551 AAATGGTGCCGGCGGTGGCCGAAGTGAGCGGGCATCTGGGGAAAAGTTG hFATP1con.seq ORF
551 AGATGGCAGCGGCGGTGCCGGAGGTGAGCGAGCAGCTGGGGAAGAGCCC mFATP1.seq ORF (from genomic)
601 ATCAAGTTCTGCTCTGGAGACTTGGGGCCCGAGGGCATCTTCCGGACAC hFATP1con.seq ORF
601 CTCAAGTTCTGCTCTGGAGATCTGGCGCCGTGAGACATCCTGCCTGACAC mFATP1.seq ORF (from genomic)
651 CCACCTGCCTGGACCCGCTGCTGAAGGAGGGCCTCTACTGCCCCCTGGCAC hFATP1con.seq ORF
651 GCAGCTCCTGGACCCCATGCTTGCTGAGGCGCCCACCACACCGCTGGCAC mFATP1.seq ORF (from genomic)
701 AGATCCCCAGCAAGGGCATGGACGATCGTCTTTCTACATCTACACGTCG hFATP1con.seq ORF
701 AAGCCCCAGGCAAGGGCATGGATGATCCGCGTTTTACATCTATACTTCT mFATP1.seq ORF (from genomic)
751 GGGACCACCGCGCCTGCCCAAGCCTGCCATTGCTCGTGCACAGCAGGTACTA hFATP1con.seq ORF
751 GGGACCACCGCGCCTTCCTAAGCCTGCCATTGCGGTGCACAGCAGGCTACTA mFATP1.seq ORF (from genomic)
801 CCGCATGGGAGCCTTCGGCCACCACGCCTACCGCATGCAGGCGGCTGACG hFATP1con.seq ORF
801 CCGCATTGCTGCCTTTGCCCACCATTCCTACAGCATGCGTCCCGCCGATG mFATP1.seq ORF (from genomic)
851 TGCTCTATGACTGCCTGCCCCTGTACCACTCGCGCAGGAAACATCATCGGC hFATP1con.seq ORF
851 TGCTCTATGACTGCCTGCCACTCTACCACTCTGCAGGGAACATCATGGGT mFATP1.seq ORF (from genomic)
901 GTGGGGCAGTGTCTCATCTATGGGCTGACAGTCGTCCTCCGCAAGAATT hFATP1con.seq ORF
901 GTGGGGCAGTGCGTCATCTACGGGTTGACGGTGGTACTGCGCAAGAAGTT mFATP1.seq ORF (from genomic)
951 CTCGGCCAGCCGCTTGTGGGACGACTGCATCAAGTACAACTGCACGGTGG hFATP1con.seq ORF
951 CTCCGCCAGCCGCTTCTGGGATGACTGTGTCAAGTACAATTGCACGGTAG mFATP1.seq ORF (from genomic)
1001 TTCAGTACATCGGGGAGATCTGCCGCTACCTGCTGAAGCAGCCGGTGCGC hFATP1con.seq ORF
1001 TGCACTACATAGCTGAAATCTGCCGCTACCTGCTGAGGCAGCCGGTTCGC mFATP1.seq ORF (from genomic)
1051 GAGGCGGAGAGGCGACACGCGCGTGCGCCTGCGGTGGGAACGGGCTGCG hFATP1con.seq ORF
1051 GACGTGGAGCAGCGACACCGCGTGCCGCCTGGGCGTGGCTAATGGGCTGG mFATP1.seq ORF (from genomic)
1101 TCCTGCCATCTGGGAGGAGTTCACGGAGCGCTTCGGCGTACGCCAAATCG hFATP1con.seq ORF
1101 GCGAGCCATCTGGCAGGAGTTCACGCAGCGCTTCGGTGTGCCACAGATCG mFATP1.seq ORF (from genomic)
1151 GGAGTTCTACGGCGCCACCGAGTGCAACTGCAGCATTGCCAACATGGAC hFATP1con.seq ORF
1151 GCGAGTTCTACGGCGCTACCGAGTGCAACTGCAGCATTGCCAACATGGAC mFATP1.seq ORF (from genomic)
1201 GGCAAGGTCGGCTCCTGTGGCTTCAACAGCCGCATCCTGCCCCACGTGTA hFATP1con.seq ORF
1201 GGCAAGCTCGGCTCCTGCGGCTTCAACAGCCGTATCCTCCACGCATGTGTA mFATP1.seq ORF (from genomic)
1251 CCCCATCCGGCTGGTGAAGGTCAATGAGGACAACAATGGAGCTGCTGCGGG hFATP1con.seq ORF
1251 CCCCATCCGTCTGCTCAAGGTCAATGAGGACAACGATGGAGCCACTGCGGC mFATP1.seq ORF (from genomic)
1301 ATGCCCAGGGCCTCTGCATCCCCTGCCAGGCCGGGGGAGCCTGGCCTCCT hFATP1con.seq ORF
1301 ACTCCGAGGGCCTCTTCATCCGGTGCCAGCCCGGGGGAACCCGGCCTTCT mFATP1.seq ORF (from genomic)

FIG. 30A

```
1351 GTGGGTCAGATCAACCAACAGGACCCGCCTGCCGCCGCTTCGATGGCCATGT hFATP1con.seq ORF
1351 GTGGGCCAGATCAACCAGCAGGACCCTCCTGCCGCCGTTTCGATCGTTATGT mFATP1.seq ORF (from genomic)

1401 CAGCGAGAGCGCCACCAGCAAGAAGATCGCCCACAGCGCTCTTCAGCAAGG hFATP1con.seq ORF
1401 TAGTGAGCAGTGCCACCAACAAGAAGATTGCCCACAGCGCTTTCCGAAAGG mFATP1.seq ORF (from genomic)

1451 GCGACAGCGCCTACCTCTCAGGTGACGTGCTAGTGATGGATGAGCTGGGC hFATP1con.seq ORF
1451 GCGATAGCGCCTACCTCTCAGGTGACGTGCTAGTGATGGACGAGCTGGGC mFATP1.seq ORF (from genomic)

1501 TACATGTACTTCCGGACCGTAGCGGGGACACCTTCCGCTGGCCAGGGGA hFATP1con.seq ORF
1501 TACATGTATTCCGTGACCGCAGCGGGACACCTTCCGCTGGCCCGGGGA mFATP1.seq ORF (from genomic)

1551 GAACGTCTCCACCAACGAGGTGGAGGGCGTGCTGAGCCGCCTGCTGGCC hFATP1con.seq ORF
1551 GAACCTGTCCACCACGGAGGTGGAAGCCGTGCTGAGCCGCCTACTGGCC mFATP1.seq ORF (from genomic)

1601 AGACAGACGTGGCCGTCTATGGGGTGGCTCTTCCAGGAGTGGAGGGTAAG hFATP1con.seq ORF
1601 ACACGGACGTGGCTGTATGCGGTGGCTCTTGCCAGGAGTGGAGGGGAAA mFATP1.seq ORF (from genomic)

1651 GCAGGGATGGGGCCCGTCGCAGACCCCACAGGCTGCTGGACCCCAAGGC hFATP1con.seq ORF
1651 GCTGGCATGGCAGCCATCGCAGATCCCCAGAGCCAGTTGGACCCTAAGT mFATP1.seq ORF (from genomic)

1701 GATATACCAGGAGCTGCAGAAGGTGCTGGCACCCTATGCCCGGCCCATCT hFATP1con.seq ORF
1701 AATGTACCAGCAATTACAGAAGGTTCTTGCATCCTATGGTCGGCCCATGT mFATP1.seq ORF (from genomic)

1751 TCCTGCGCCTCCGTGCCCAGGTGGACCACCACAGGCACCTTCAAGATGCAG hFATP1con.seq ORF
1751 TCCTGCGTCTTCTGCCCCAGGTGGATACCACAGGCACCTTCAAGATCCAG mFATP1.seq ORF (from genomic)

1801 AAGACGAGGCTGCAGCGAGAGGGCTTTGACCCACGCCAGACCTCAGACCG hFATP1con.seq ORF
1801 AAGACCCGGCTGCAGCGTGAAGGCTTTCACCCCGTCAGACCTCACACAG mFATP1.seq ORF (from genomic)

1851 GCTGTCCTTCCTGGACGTGAAGCAGCGCCACTACCTGCCGTTAAATGAGG hFATP1con.seq ORF
1851 GCTCTTCTTTCTAGACCTGAAGCAGGGACGCTATGTACCCCTGGATGAGA mFATP1.seq ORF (from genomic)

1901 CAGTGTACAGTCGCATCTGCTCGGGCGCCTTCGCCCTCTGA hFATP1con.seq ORF
1901 GAGTGCATGCCGCATTTGTGCAGGCGACTTGTCACTG mFATP1.seq ORF (from genomic)
```

Decoration 'Decoration #1': Shade (with solid bright yellow) residues that match the consensus named 'Consensus #1' exactly.

FIG. 30B

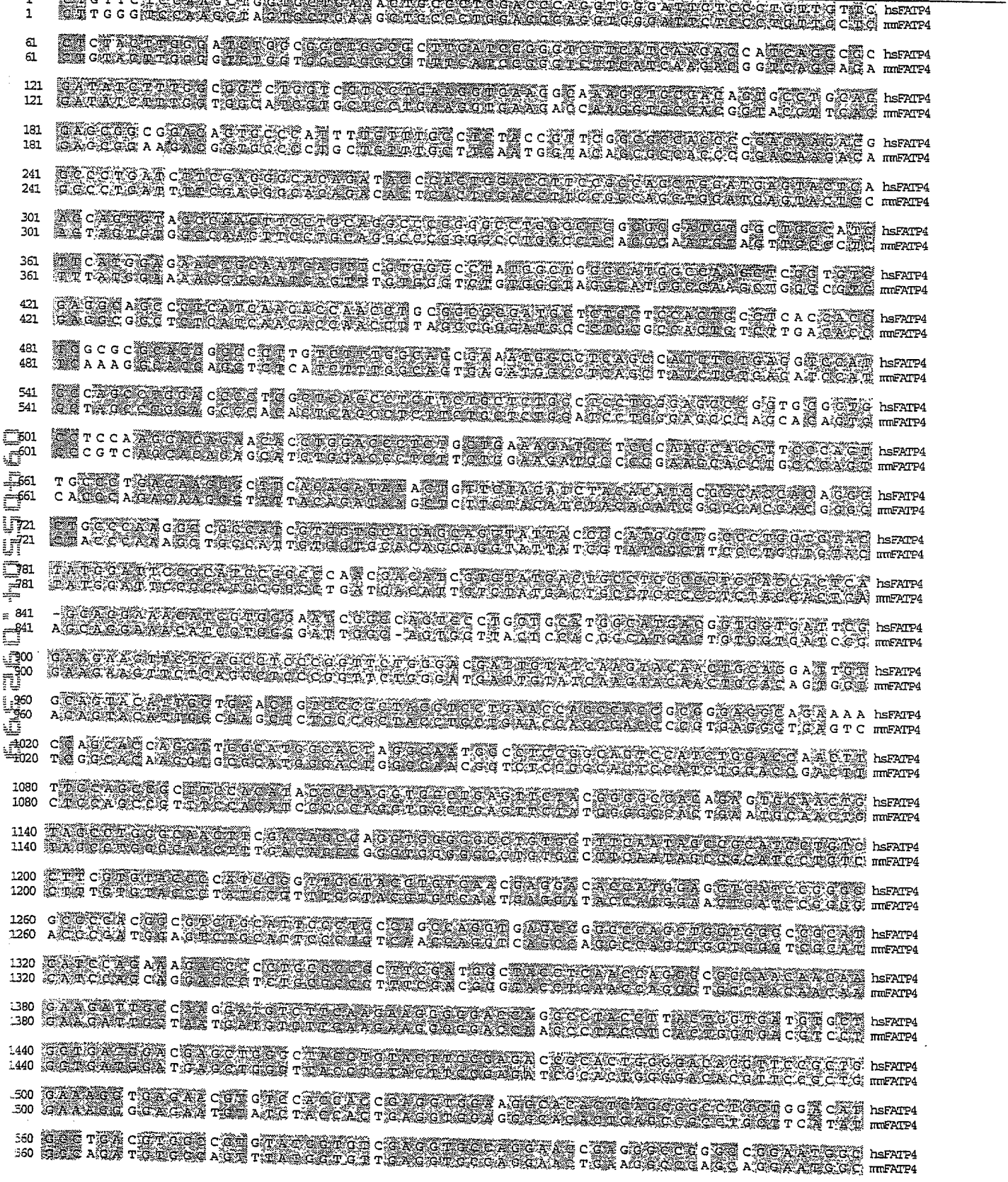
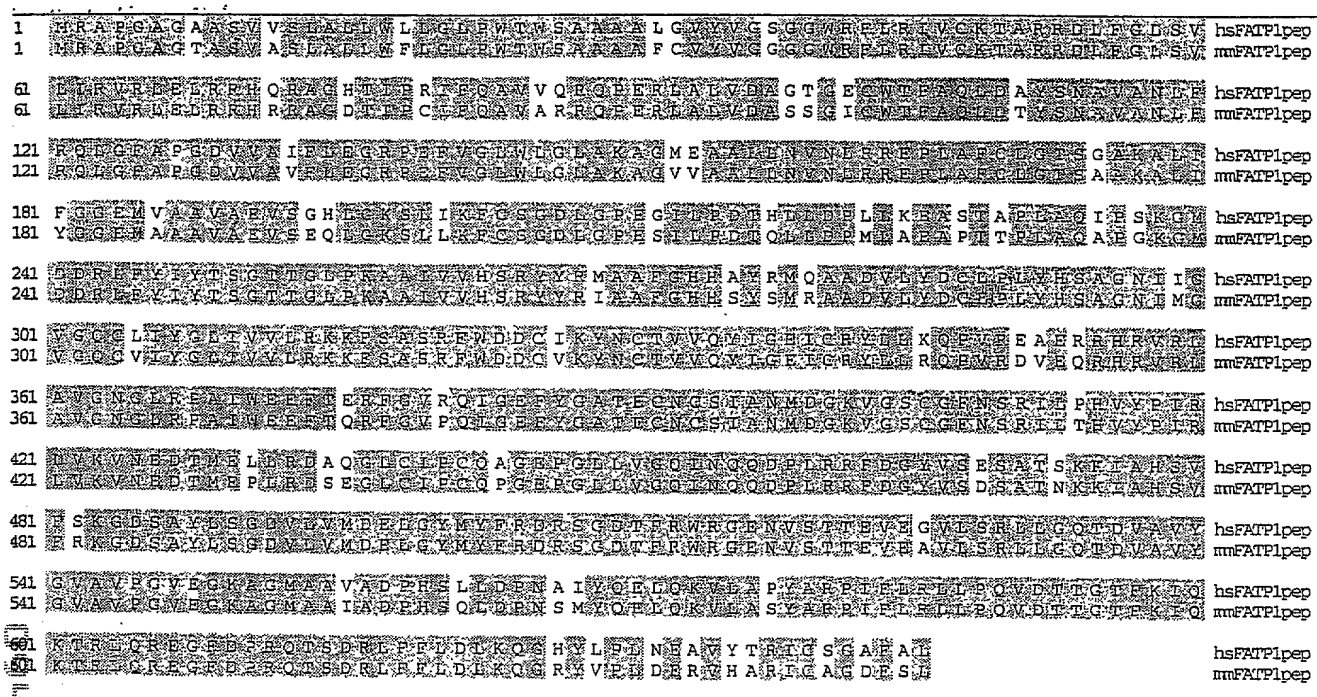


Figure 3/A

```
1620 TGCTGTGGCCAGCCCCACTGGCAACTGTGACCTGGAGCGCTTTGCTCAGGTCTTGGAGAA hsFATP4
1620 TGGCGTTGCAAGTGCC-TCAGCAACTGTGACCTGGAGAGCTTTGCACAGACCTTCAAAAA mmFATP4

1680 GCAACTGCCCCTGTATGCGGCCCCCATCTTCCTGCGCCTCCTGCCTGAGCTGCACAAAAC hsFATP4
1680 GGAGCTGCCTCTGTATCCCCCGCCCATCTTCCTCCGCTTCTTGCCTGAGCTGCACAAGAC mmFATP4

1740 AGGAACCTACAAGTTCCAGAAGACAGAGCTACGGAAGGAGGGCTTTGACCCGGCTATTGT hsFATP4
1740 AGGGACCTTCAAGTTCCAGAAGACAGAGTTGCGGAAGGAGGGCTTTGACCCATCTGTTGT mmFATP4

1800 GAAAGACCCGCTGTTCTATCTAGATGCCCAGAAGGGCCGCTACGTCCCGCTGCACCAAGA hsFATP4
1800 GAAAGACCCGCTGTTCTATCTGCATGCTCGGAAGGGCTGCTACGTTGCACTGGACCAGGA mmFATP4

1860 GGCCTACAGCCGCATCCAGGCAGGCGAGGAGAAGCTG hsFATP4
1860 GGCCTATACCCGCATCCAGGCAGGCGAGGAGAAGCTG mmFATP4
```

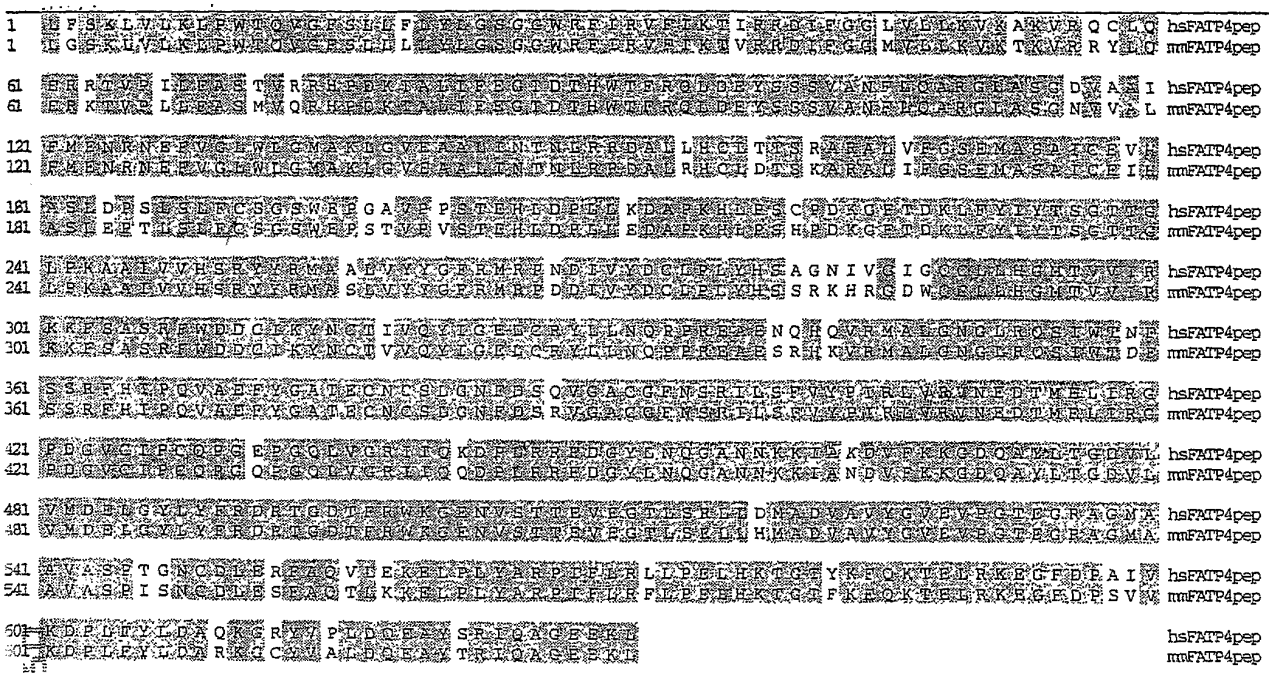
Decoration 'Decoration #1': Shade (with solid bright yellow) residues that match the consensus named 'Consensus #1' exactly.

Figure 31B



Decoration 'Decoration #2': Shade (with solid bright yellow) residues that match the consensus named 'Consensus #1' exactly.

Figure 32



Decoration 'Decoration #1': Shade (with solid bright yellow) residues that match the consensus named 'Consensus #1' exactly.

Figure 33

hsFATP6

```
1       aac ggc aag taa gcg caa cgc aat taa tgt gag tag ctc act cat tag gca ccc cag gct
61      tta cac ttt atg ctt ccg ggc tcg tat gtt gtg tgg aat tgt gag cgg ata cca att tca
121     cac agg aac cag cta tga cat gat tac gaa ttt aat acg act cac tat agg gaa ttt ggc
181     cct cga ggc caa gaa ttc ggc acg agg ggt gct gag ccc ctg cgc ggt ttc tgg tgc gta
241     gag act gta aat cgc tgc gct tct cag tca tca tca tcc cag ctt ttc ccg gct cga att
301     cag cct cca act caa gct cgc ggg aaa gac tac ctg aga gga gaa aag ctt ctg tcc ctg
361     gac ctt ctt ctg agg gtg gag tcg gag gct ccc tgc ttt cca gcc gcc cag tga ccc aag
421     ctt aat ctt cag cac cac ttg ggg cga cct ttt cgg tgc aaa cct acg att ctg ttt ctc
481     agg act cct ccc cat ccc gct tcg ccc cgg aaa agc tga caa gaa ctt cag gtg taa gcc
541     ctg agt agt gag gat ctg cgg tct ccg tgg aga gct gtg cct gga aga gaa gga cgc tgg
601     tgg ggg ctg aga tca gag ctg tct tct ggc cca gtt gcc ccc atg ctt ctg tca tgg cta
                                                        M   L   L   S   W   L
661     aca gtt cta ggg gct gga atg gtc gtc ctg cac ttc ttg cag aaa ctc ctg ttc cct tac
        T   V   L   G   A   G   M   V   V   L   H   F   L   Q   K   L   L   F   P   Y
721     ttt tgg gat gac ttc tgg ttc gtg ttg aag gtg gtg ctc att ata att cgg ctg aag aag
        F   W   D   D   F   W   F   V   L   K   V   V   L   I   I   I   R   L   K   K
781     tat gaa aag aga ggg gag ctg gtg act gtg ctg gat aaa ttc ttg agt cat gcc aaa aga
        Y   E   K   R   G   E   L   V   T   V   L   D   K   F   L   S   H   A   K   R
841     caa cct cgg aaa cct ttc atc atc tat gag gga gac atc tac acc tat cag gat gta gac
        Q   P   R   K   P   F   I   I   Y   E   G   D   I   Y   T   Y   Q   D   V   D
901     aaa agg agc agc aga gtg gcc cat gtc ttc ctg aac cat tcc tct ctg aaa aag ggg gac
        K   R   S   S   R   V   A   H   V   F   L   N   H   S   S   L   K   K   G   D
961     acg gtg gct ctg ctg atg agc aat gag ccg gac ttc gtt cac gtg tgg ttc ggc ctc gcc
        T   V   A   L   L   M   S   N   E   P   D   F   V   H   V   W   F   G   L   A
1021    aag ctg ggc tgc gtg gtg gcc ttt ctc aac acc aac att cgc tcc aac tcc ctc ctg aat
        K   L   G   C   V   V   A   F   L   N   T   N   I   R   S   N   S   L   L   N
1081    tgc atc cgc gcc tgt ggg ccc aga gcc cta gtg gtg ggc gca gat ttg ctt gga acg gta
        C   I   R   A   C   G   P   R   A   L   V   V   G   A   D   L   L   G   T   V
1141    gaa gaa atc ctt cca agc ctc tca gaa aat atc agt gtt tgg ggg atg aaa gat tct gtt
        E   E   I   L   P   S   L   S   E   N   I   S   V   W   G   M   K   D   S   V
1201    cca caa ggt gta att tca ctc aaa gaa aaa ctg agc acc tca cct gat gag ccc gtg cca
        P   Q   G   V   I   S   L   K   E   K   L   S   T   S   P   D   E   P   V   P
1261    cgc agc cac cat gtt gtc tca ctc ctc aag tct act tgt ctt tac att ttt acc tct gga
        R   S   H   H   V   V   S   L   L   K   S   T   C   L   Y   I   F   T   S   G
1321    aca aca ggt cta cca aaa gca gct gtg att agt cag ctg cag gtt tta agg ggt tct gct
        T   T   G   L   P   K   A   A   V   I   S   Q   L   Q   V   L   R   G   S   A
1381    gtc ctg tgg gct ttt ggt tgt act gct cat gac att gtt tat ata acc ctt cct ctg tat
        V   L   W   A   F   G   C   T   A   H   D   I   V   Y   I   T   L   P   L   Y
1441    cat agt tca gca gct atc ctg gga att tct gga tgt gtt gag ttg ggt gcc act tgt gtg
        H   S   S   A   A   I   L   G   I   S   G   C   V   E   L   G   A   T   C   V
1501    tta aag aag aaa ttt tca gca agc cag ttt tgg agt gac tgc aag aag tat gat gtg act
        L   K   K   K   F   S   A   S   Q   F   W   S   D   C   K   K   Y   D   V   T
1561    gtg ttt cag tat att gga gaa ctt tgt cgc tac ctt tgc aaa caa tct aag aga gaa gga
        V   F   Q   Y   I   G   E   L   C   R   Y   L   C   K   Q   S   K   R   E   G
1621    gaa aag gat cat aag gtg cgt ttg gca att gga aat ggc ata cgg agt gat gta tgg aga
        E   K   D   H   K   V   R   L   A   I   G   N   G   I   R   S   D   V   W   R
1681    gaa ttt tta gac aga ttt gga aat ata aag gtg tgt gaa ctt tat gca gct acc gaa tca
        E   F   L   D   R   F   G   N   I   K   V   C   E   L   Y   A   A   T   E   S
1741    agc ata tct ttc atg aac tac act ggg aga att gga gca att ggg aga aca aat ttg ttt
        S   I   S   F   M   N   Y   T   G   R   I   G   A   I   G   R   T   N   L   F
1801    tac aaa ctt ctt tcc act ttt gac tta ata aag tat gac ttt cag aaa gat gaa ccc atg
        Y   K   L   L   S   T   F   D   L   I   K   Y   D   F   Q   K   D   E   P   M
1861    aga aat gag cag ggt tgg tgt att cat gtg aaa aaa gga gaa cct gga ctt ctc att tct
        R   N   E   Q   G   W   C   I   H   V   K   K   G   E   P   G   L   L   I   S
1921    cga gtg aat gca aaa aat ccc ttc ttt ggc tat gct ggg cct tat aag cac aca aaa gac
        R   V   N   A   K   N   P   F   F   G   Y   A   G   P   Y   K   H   T   K   D
1981    aaa ttg ctt tgt gat gtt ttt aag aag gga gat gtt tac ctt aat act gga gac tta ata
        K   L   L   C   D   V   F   K   K   G   D   V   Y   L   N   T   G   D   L   I
2041    gtc cag gat cag gac aat ttc ctt tat ttt tgg gac cgt act gga gac act ttc aga tgg
        V   Q   D   Q   D   N   F   L   Y   F   W   D   R   T   G   D   T   F   R   W
2101    aaa gga gaa aat gtc gca acc act gag gtt gct gat gtt att gga atg ttg gat ttc ata
        K   G   E   N   V   A   T   T   E   V   A   D   V   I   G   M   L   D   F   I
2161    cag gaa gca aac gtc tat ggt gtg gct ata tca ggt tat gaa gga aga gca gga atg gct
        Q   E   A   N   V   Y   G   V   A   I   S   G   Y   E   G   R   A   G   M   A
2221    tct att att tta aaa cca aat aca tct tta gat ttg gaa aaa gtt tat gaa caa gtt gta
        S   I   I   L   K   P   N   T   S   L   D   L   E   K   V   Y   E   Q   V   V
2281    aca ttt cta cca gct tat gct tgt cca cga ttt tta aga att cag gaa aaa atg gaa gca
        T   F   L   P   A   Y   A   C   P   R   F   L   R   I   Q   E   K   M   E   A
2341    aca gga aca ttc aaa cta ttg aag cat cag ttg gtg gaa gat gga ttt aat cca ctg aaa
        T   G   T   F   K   L   L   K   H   Q   L   V   E   D   G   F   N   P   L   K
2401    att tct gaa cca ctt tac ttc atg gat aac ttg aaa aag tct tat gtt cta ctg acc agg
        I   S   E   P   L   Y   F   M   D   N   L   K   K   S   Y   V   L   L   T   R
2461    gaa ctt tat gat caa ata atg tta ggg gaa ata aaa ctt taa gat ttt tat atc tag aac
        E   L   Y   D   Q   I   M   L   G   E   I   K   L   *
2521    ttt cat atg ctt tct tag gaa gag tga gag ggg ggt ata tga ttc ttt atg aaa tgg gga
2581    aag gga gct aac att aat tat gca tgt act ata ttt cct taa tat gag aga taa ttt ttt
2641    aat tgc ata aga att tta att tct ttt aat tga tat aaa cat tag ttg att att ctt ttt
2701    atc tat ttg gag att cag tgc ata act aag tat ttt cct taa tac taa aga ttt taa ata
2761    ata aat agc ggc tag cgg ttt gga caa tca cta aaa atg tac ttt cta ata agt aaa att
2821    tct aat ttt gaa taa aag att aaa ttt tac tga aaa aaa aaa aaa aaa aaa aaa ttg gcg
2881    gcc gc
```

Figure 34

Protein sequence 619 a.a. MLLSWLTVLGAG ... LYDQIMLGEIKL

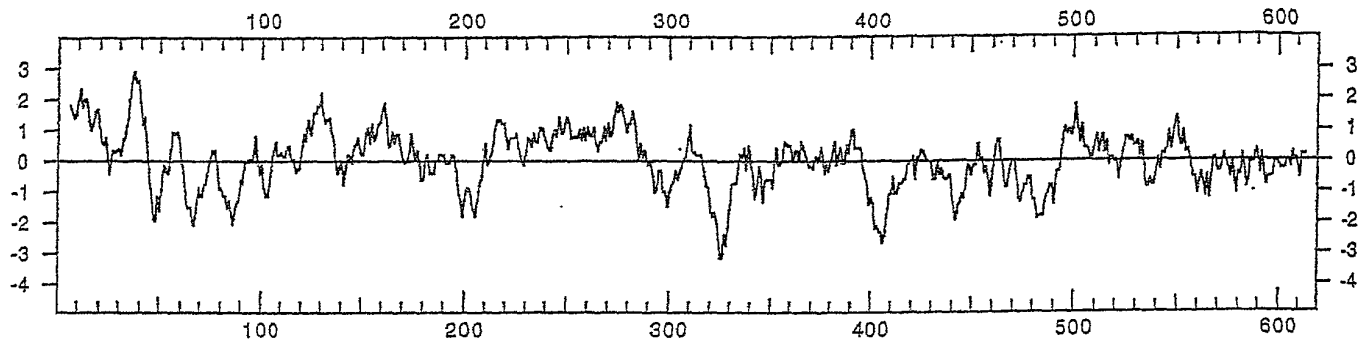


FIGURE 35 A

Protein sequence 619 a.a. MLLSWLTVLGAG ... LYDQIMLGEIKL

619 Amino Acids MW : 70066 Dalton

			n	n(%)	MW	MW(%)
A	ala	alanine	33	5.3	2344	3.3
C	cys	cysteine	14	2.3	1442	2.1
D	asp	aspartic acid	34	5.5	3910	5.6
E	glu	glutamic acid	31	5.0	4000	5.7
F	phe	phenylalanine	34	5.5	5000	7.1
G	gly	glycine	44	7.1	2508	3.6
H	his	histidine	13	2.1	1781	2.5
I	ile	isoleucine	37	6.0	4184	6.0
K	lys	lysine	48	7.8	6148	8.8
L	leu	leucine	75	12.1	8481	12.1
M	met	methionine	11	1.8	1441	2.1
N	asn	asparagine	21	3.4	2394	3.4
P	pro	proline	21	3.4	2038	2.9
Q	gln	glutamine	18	2.9	2305	3.3
R	arg	arginine	27	4.4	4214	6.0
S	ser	serine	40	6.5	3481	5.0
T	thr	threonine	30	4.8	3031	4.3
V	val	valine	51	8.2	5052	7.2
W	trp	tryptophan	11	1.8	2046	2.9
X	ukw	unknown	-	-		
Y	tyr	tyrosine	26	4.2	4239	6.1
Z	---	STOP	-	-		

FIGURE 35 B

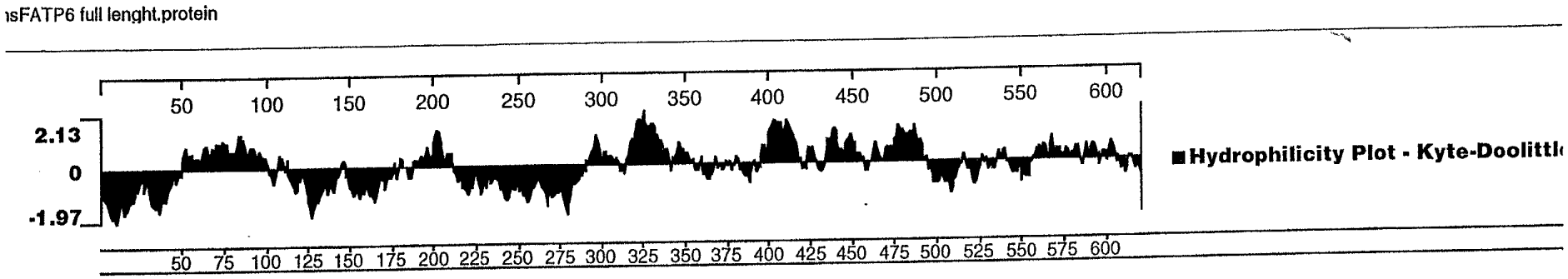
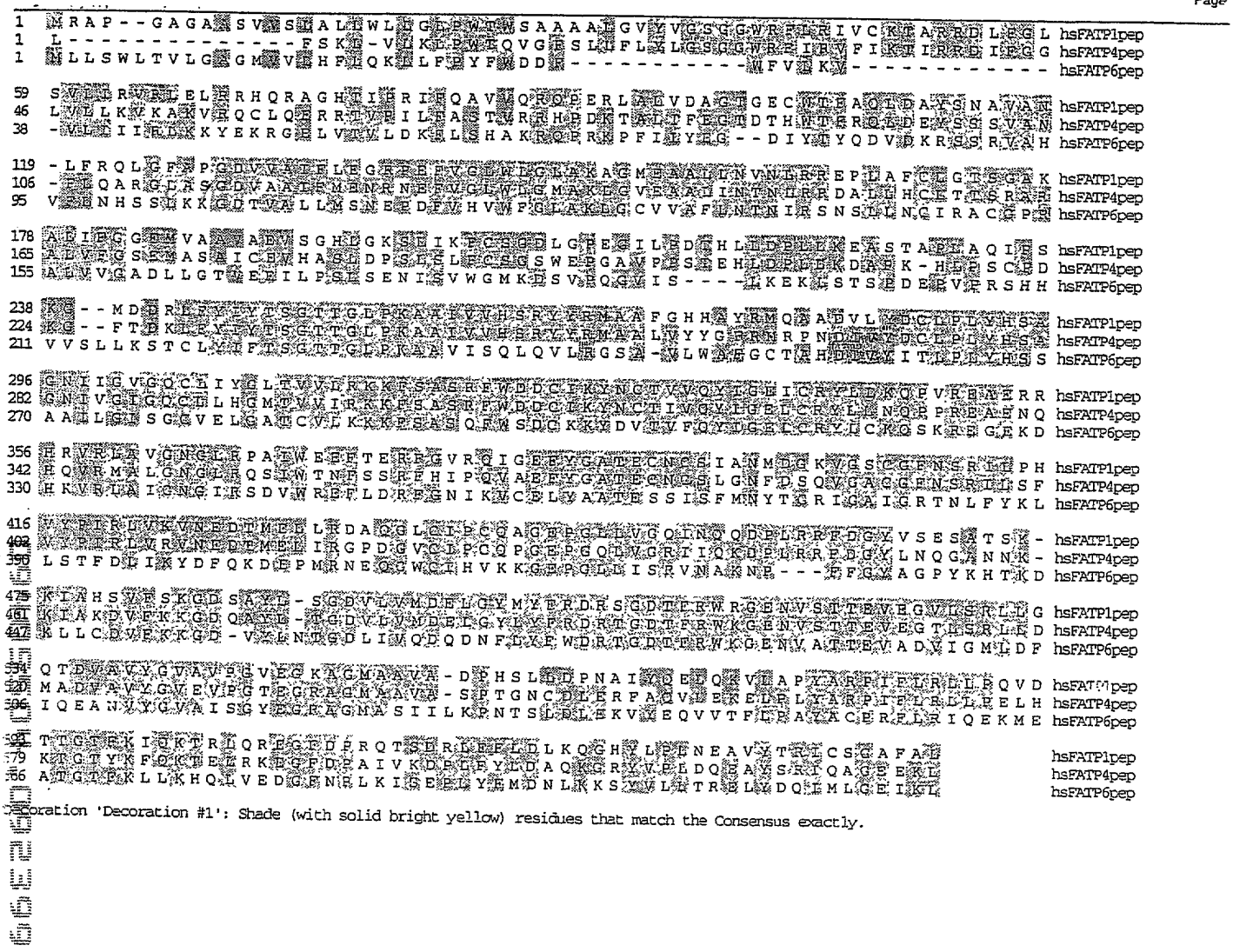


Figure 35C



Decoration 'Decoration #1': Shade (with solid bright yellow) residues that match the Consensus exactly.

Figure 36

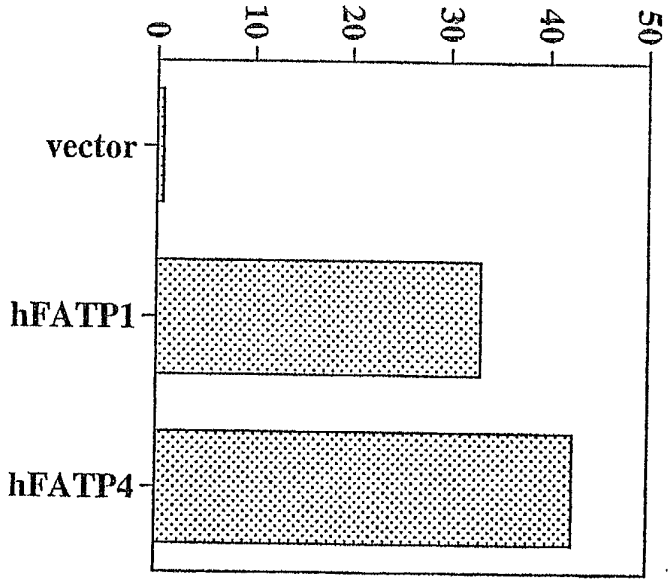
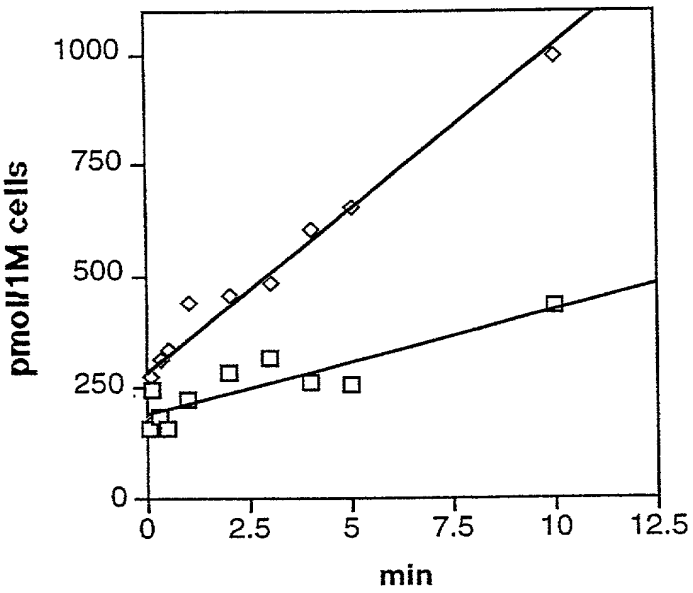


Figure 37



□ 293 vector control: 23 pmol/(min*1*10^6 cells)

◇ 293 FATP4 clone 7: 73 pmol/(min*1*10^6 cells)

Fig. 38

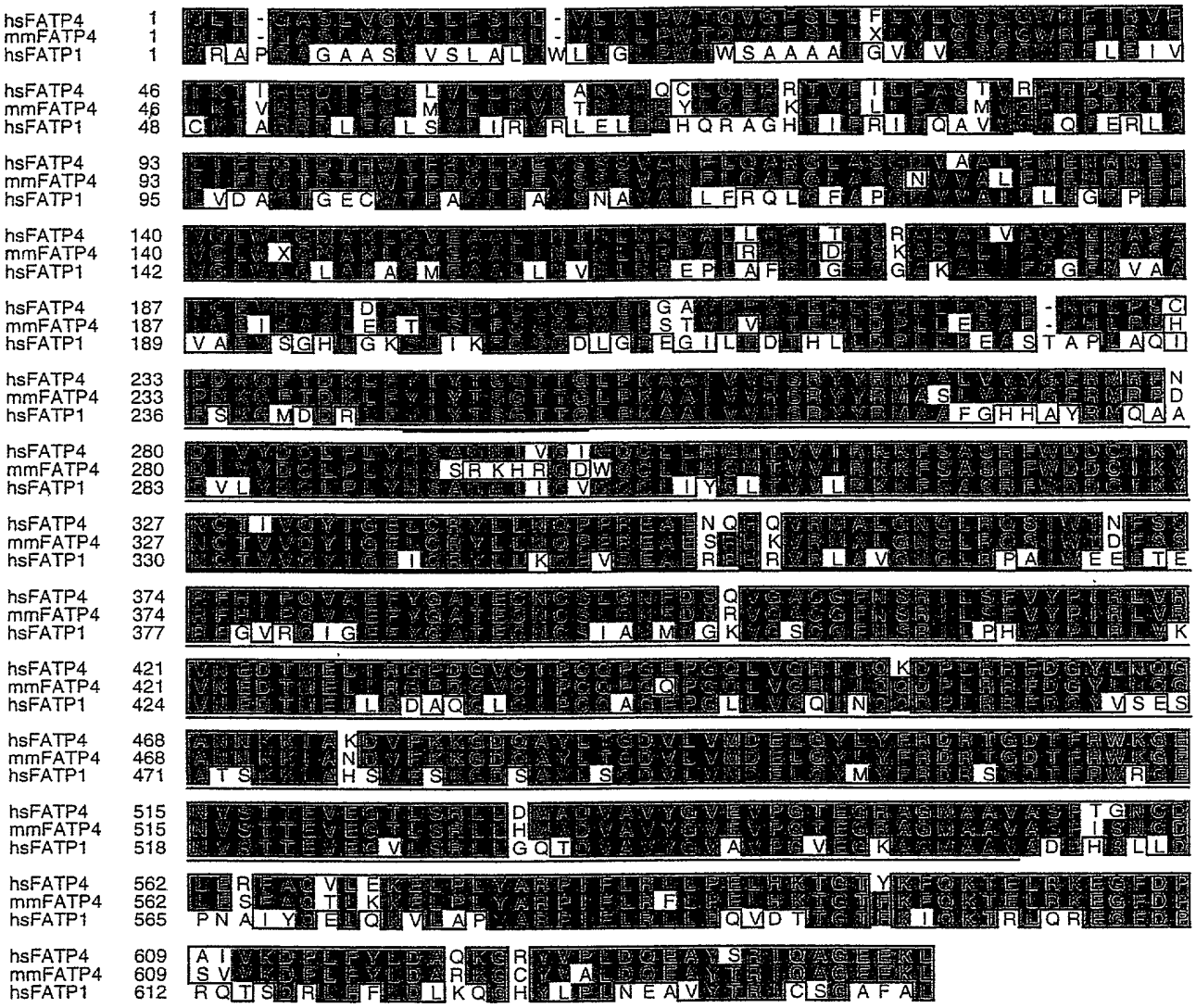
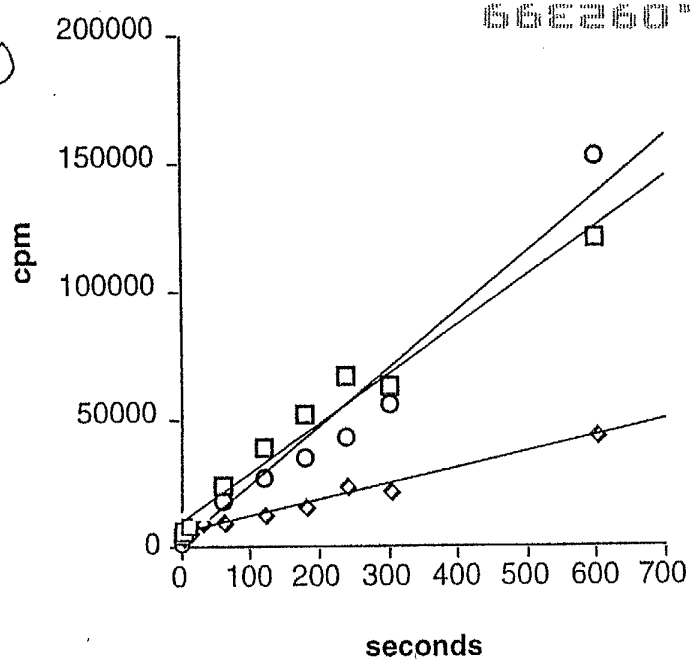

hsFATP4 1
mmFATP4 1
hsFATP1 1
hsFATP4 46
mmFATP4 46
hsFATP1 48
hsFATP4 93
mmFATP4 93
hsFATP1 95
hsFATP4 140
mmFATP4 140
hsFATP1 142
hsFATP4 187
mmFATP4 187
hsFATP1 189
hsFATP4 233
mmFATP4 233
hsFATP1 236
hsFATP4 280
mmFATP4 280
hsFATP1 283
hsFATP4 327
mmFATP4 327
hsFATP1 330
hsFATP4 374
mmFATP4 374
hsFATP1 377
hsFATP4 421
mmFATP4 421
hsFATP1 424
hsFATP4 468
mmFATP4 468
hsFATP1 471
hsFATP4 515
mmFATP4 515
hsFATP1 518
hsFATP4 562
mmFATP4 562
hsFATP1 565
hsFATP4 609
mmFATP4 609
hsFATP1 612

Fig. 39

Fig. 40



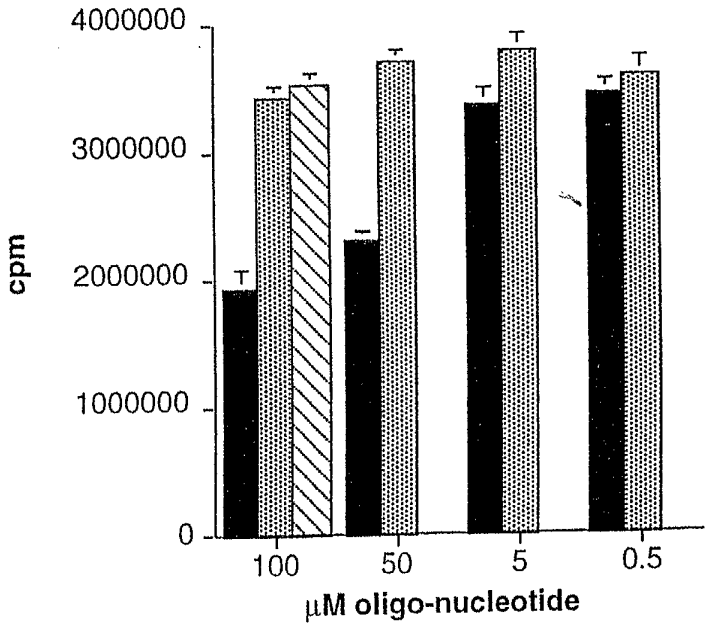
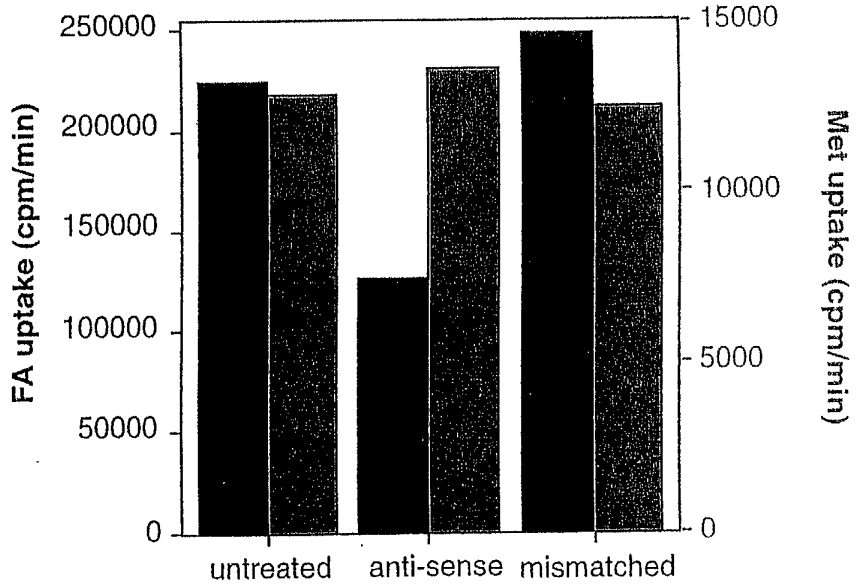


Fig. 41

Fig. 42



mmFATP4 DNA sequence

ATGCTGCTTGGAGCCTCTCTGGTGGGGGCGCTACTGTTCTCCAAGCTAGTGCTGAAGCTGCCCTGGACCCAGGTGGGATT
CTCCCTGTTGCTCCTGTACTTGGGGTCTGGTGGCTGGCGTTTCATCCGGGTCTTCATCAAGACGGTCAGGAGAGATATCT
TTGGTGGCATGGTGCTCCTGAAGGTGAAGACCAAGGTGCGACGGTACCTTCAGGAGCGGAAGACGGTGCCCCTGCTGTTT
GCTTCAATGGTACAGCGCCACCCGGACAAGACAGCCCTGATTTTCGAGGGCACAGACACTCACTGGACCTTCCGCCAGCT
GGATGAGTACTCCAGTAGTGTGGCCAACTTCCTGCAGGCCCGGGGCCTGGCCTCAGGCAATGTAGTTGCCCTCTTTATGG
AAAACCGCAATGAGTTTGTGGGTCTGTGGCTAGGCATGGCCAAGCTGGGCGTGGAGGCGGCTCTCATCAACACCAACCTT
AGGCGGGATGCCCTGCGCCACTGTCTTGACACCTCAAAGGCACGAGCTCTCATCTTTGGCAGTGAGATGGCCTCAGCTAT
CTGTGAGATCCATGCTAGCCTGGAGCCCACACTCAGCCTCTTCTGCTCTGGATCCTGGGAGCCCAGCACAGTGCCCGTCA
GCACAGAGCATCTGGACCCTCTTCTGGAAGATGCCCCGAAGCACCTGCCCAGTCACCCAGACAAGGGTTTTACAGATAAG
CTCTTCTACATCTACACATCGGGCACCACGGGGCTACCCAAAGCTGCCATTGTGGTGCACAGCAGGTATTATCGTATGGC
TTCCCTGGTGTACTATGGATTCCGCATGCGGCCTGATGACATTGTCTATGACTGCCTCCCCCTCTACCACTCAAGCAGGA
AACATCGTGGGGATTGGCAGTGCTTACTCCACGGCATGACTGTGGTGATCCGGAAGAAGTTCTCAGCCTCCCGGTTCTGG
GATGATTGTATCAAGTACAACTGCACAGTGGTACAGTACATTGGCGAGCTCTGCCGCTACCTCCTGAACCAGCCACCCCG
TGAGGCTGAGTCTCGGCACAAGGTGCGCATGGCACTGGGCAACGGTCTCCGGCAGTCCATCTGGACCGACTTCTCCAGCC
GTTTCCACATCCCCCAGGTGGCTGAGTTCTATGGGGCCACTGAATGCAACTGTAGCCTGGGCAACTTTGACAGCCGGGTG
GGGGCCTGTGGCTTCAATAGCCGCATCCTGTCCTTTGTGTACCCTATCCGTTTGGTACGTGTCAATGAGGATACCATGGA
ACTGATCCGGGGACCCGATGGAGTCTGCATTCCCTGTCAACCAGGTCAGCCAGGCCAGCTGGTGGGTCGCATCATCCAGC
AGGACCCTCTGCGCCGTTTCGACGGGTACCTCAACCAGGGTGCCAACAACAAGAAGATTGCTAATGATGTCTTCAAGAAG
GGGGACCAAGCCTACCTCACTGGTGACGTCCTGGTGATGGATGAGCTGGGTTACCTGTACTTCCGAGATCGCACTGGGGA
CACGTTCCGCTGGAAAGGGGAGAATGTATCTACCACTGAGGTGGAGGGCACACTCAGCCGCCTGCTTCATATGGCAGATG
TGGCAGTTTATGGTGTTGAGGTGCCAGGAACTGAAGGCCGAGCAGGAATGGCTGCCGTTGCAAGTCCCATCAGCAACTGT
GACCTGGAGAGCTTTGCACAGACCTTGAAAAAGGAGCTGCCTCTGTATGCCCGCCCCATCTTCCTGCGCTTCTTGCCTGA
GCTGCACAAGACAGGGACCTTCAAGTTCCAGAAGACAGAGTTGCGGAAGGAGGGCTTTGACCCATCTGTTGTGAAAGACC
CGCTGTTCTATCTGGATGCTCGGAAGGGCTGCTACGTTGCACTGGACCAGGAGGCCTATACCCGCATCCAGGCAGGCGAG
GAGAAGCTGTGATTTCCCCCTACATCCCTCTGAGGGCCAGAAGATGCTGGATTCAGAGCCCTAGCGTCCACCCCAGAGGG
TCCTGGGCAATGCCAGACCAAAGCTAGCAGGGCCCGCACCTCCGCCCCTAGGTGCTGATCTCCCCTCTCCCAAACTGCCA
AGTGACTCACTGCCGCTTCCCCGACCCTCCAGAGGCTTTCTGTGAAAGTCTCATCCAAGCTGTGTCTTCTGGTCCAGGCG
TGGCCCCTGGCCCCAGGGTTTCTGATAGGCTCCTTTAGGATGGTATCTTGGGTCCAGCGGGCCAGGGTGTGGGAGAGGAG
TCACTAAGATCCCTCCAATCAGAAGGGAGCTTACAAAGGAACCAAGGCAAAGCCTGTAGACTCAGGAAGCTAAGTGGCCA
GAGACTATAGTGGCCAGTCATCCCATGTCCACAGAGGATCTTGGTCCAGAGCTGCCAAAGTGTCACCTCTCCCTGCCTGC

ACCTCTGGGGAAAAGAGGACAGCATGTGGCCACTGGGCACCTGTCTCAAGAAGTCAGGATCACACACTCAGTCCTTGTTT
CTCCAGGTTCCCTTGTTCTTGTCTCGGGGAGGGAGGGACGAGTGTCCTGTCTGTCCTTCCTGCCTGTCTGTGAGTCTGTG
TTGCTTCTCCATCTGTCCTAGCCTGAGTGTGGGTGGAACAGGCATGAGGAGAGTGTGGCTCAGGGGCCAATAAACTCTGC
CTTGACTCCTCTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

Figure 43A

mmFATP4 protein sequence

MLLGASLVGALLFSKLVLKLPWTQVGFSLLLLYLGSGGWRFIRVFIKTVRRDIFGGMVLLKVKTKVRRYLQERKTVPLLF
ASMVQRHPDKTALIFEGTDTHWTFRQLDEYSSSVANFLQARGLASGNVVALFMENRNEFVGLWLGMAKLGVEAALINTNL
RRDALRHCLDTSKARALIFGSEMASAICEIHASLEPTLSLFCSGSWEPSTVPVSTEHLDPLLEDAPKHLPSHPDKGFTDK
LFYIYTSGTTGLPKAAIVVHSRYYRMASLVYYGFRMRPDDIVYDCLPLYHSSRKHRGDWQCLLHGMTVVIRKKFSASRFW
DDCIKYNCTVVQYIGELCRYLLNQPPREAESRHKVRMALGNGLRQSIWTDFSSRFHIPQVAEFYGATECNCSLGNFDSRV
GACGFNSRILSFVYPIRLVRVNEDTMELIRGPDGVCIPCQPGQPGQLVGRIIQQDPLRRFDGYLNQGANNKKIANDVFKK
GDQAYLTGDVLVMDELGYLYFRDRTGDTFRWKGENVSTTEVEGTLSRLLHMADVAVYGVEVPGTEGRAGMAAVASPISNC
DLESFAQTLKKELPLYARPIFLRFLPELHKTGTFKFQKTELRKEGFDPSVVKDPLFYLDARKGCYVALDQEAYTRIQAGE
EKL

Figure 43B

hsFATP1 full lenght.DNA

```
        10        20        30        40
TCGACCCACGGCGTCCGGGACCCCAAAGCAGAAGCCCGCA 40
CAGTAGGCACAGCGCACCCAAGAAGGGTCCAGGAGTCTGC 80
AGAAACAGAAAGGTCCCCGGCCTCAGCCTCCTAGTCCCTG 120
CCTGCCTCCTGCCTGAGCTTCTGGGAGACTGAAGGCACGG 160
CTTGCAGCTTCAGGATGCGGGCTCCGGGTGCGGGCGCGGC 200
       210       220       230       240
CTCGGTGGTCTCGCTGGCGCTGTTGTGGCTGCTGGGGCTG 240
CCGTGGACCTGGAGCGCGGCAGCGGCGCTCGGCGTGTACG 280
TGGCAGCGGCGGCTGGCGCTTCCTGCGCATCGTCTGCAA 320
GACCGCGAGGCGAGACCTCTTCGGTCTCTCTGTGCTGATC 360
CGCGTGCGCCTGGAGCTGCGGCGGCACCAGCGTGCCGGCC 400
       410       420       430       440
ACACCATCCCGCGCATCTTTCAGGCGGTAGTGCAGCGACA 440
GCCCGAGCGCCTGGCGCTGGTGGATGCCGGGACCGGCGAG 480
TGCTGGACCTTTGCGCAGCTGGACGCCTACTCCAATGCGG 520
TAGCCAACCTCTTCCGCCAGCTGGGCTTCGCGCCGGGCGA 560
CGTGGTGGCCATCTTCCTGGAGGGCCGGCCGGAGTTCGTG 600
       610       620       630       640
GGGCTGTGGCTGGGCCTGGCCAAGGCGGGCATGGAGGCCG 640
CGCTGCTCAACGTGAACCTGCGGCGCGAGCCCCTGGCCTT 680
CTGCCTGGGCACCTCGGGCGCTAAGGCCCTGATCTTTGGA 720
GGAGAAATGGTGGCGGCGGTGGCCGAAGTGAGCGGGCATC 760
TGGGGAAAAGTTTGATCAAGTTCTGCTCTGGAGACTTGGG 800
       810       820       830       840
GCCCGAGGGCATCTTGCCGGACACCCACCTCCTGGACCCG 840
CTGCTGAAGGAGGCCTCTACTGCCCCCTTGGCACAGATCC 880
CCAGCAAGGGCATGGACGATCGTCTTTTCTACATCTACAC 920
GTCGGGGACCACCGGGCTGCCCAAGGCTGCCATTGTCGTG 960
CACAGCAGGTACTACCGCATGGCAGCCTTCGGCCACCACG 1000
      1010      1020      1030      1040
CCTACCGCATGCAGGCGGCTGACGTGCTCTATGACTGCCT 1040
GCCCCTGTACCACTCGGCAGGAAACATCATCGGCGTGGGG 1080
CAGTGTCTCATCTATGGGCTGACAGTCGTCCTCCGCAAGA 1120
AATTCTCGGCCAGCCGCTTCTGGGACGACTGCATCAAGTA 1160
CAACTGCACGGTGGTTCAGTACATCGGGGAGATCTGCCGC 1200
```

Fig. 44A

hsFATP1 full lenght.DNA

```
        1210        1220        1230        1240
TACCTGCTGAAGCAGCCGGTGCGCGAGGCGGAGAGGCGAC 1240
ACCGCGTGCGCCTGGCGGTGGGGAACGGGCTGCGTCCTGC 1280
CATCTGGGAGGAGTTCACGGAGCGCTTCGGCGTACGCCAA 1320
ATCGGGGAGTTCTACGGCGCCACCGAGTGCAACTGCAGCA 1360
TTGCCAACATGGACGGCAAGGTCGGCTCCTGTGGTTTCAA 1400
        1410        1420        1430        1440
CAGCCGCATCCTGCCCCACGTGTACCCCATCCGGCTGGTG 1440
AAGGTCAATGAGGACACAATGGAGCTGCTGCGGGATGCCC 1480
AGGGCCTCTGCATCCCCTGCCAGGCCGGGGAGCCTGGCCT 1520
CCTTGTGGGTCAGATCAACCAACAGGACCCGCTGCGCCGC 1560
TTCGATGGCTATGTCAGCGAGAGCGCCACCAGCAAGAAGA 1600
        1610        1620        1630        1640
TCGCCCACAGCGTCTTCAGCAAGGGCGACAGCGCCTACCT 1640
CTCAGGTGACGTGCTAGTGATGGATGAGCTGGGCTACATG 1680
TACTTCCGGGACCGTAGCGGGGACACCTTCCGCTGGCGAG 1720
GGGAGAACGTCTCCACCACCGAGGTGGAGGGCGTGCTGAG 1760
CCGCCTGCTGGGCCAGACAGACGTGGCCGTCTATGGGGTG 1800
        1810        1820        1830        1840
GCTGTTCCAGGAGTGGAGGGTAAGGCAGGGATGGCGGCCG 1840
TCGCAGACCCCCACAGCCTGCTGGACCCCAACGCGATATA 1880
CCAGGAGCTGCAGAAGGTGCTGGCACCCTATGCCCGGCCC 1920
ATCTTCCTGCGCCTCCTGCCCCAGGTGGACACCACAGGCA 1960
CCTTCAAGATCCAGAAGACGAGGCTGCAGCGAGAGGGCTT 2000
        2010        2020        2030        2040
TGACCCACGCCAGACCTCAGACCGGCTCTTCTTCCTGGAC 2040
CTGAAGCAGGGCCACTACCTGCCCTTAAATGAGGCAGTCT 2080
ACACTCGCATCTGCTCGGGCGCCTTCGCCCTCTGAAGCTG 2120
TTCCTCTACTGGCCACAAACTCTGGGCCTGGTGGGAGAGG 2160
CCAGCTTGAGCCAGACAGCGCTGCCCAGGGGTGGCCGCCT 2200
        2210        2220        2230        2240
AGTACACACCCACCTGGCCGAGCTGTACCTGGCACGGCCC 2240
ATCCTGGACTGAGAAACTGGAACCTCAGAGGAACCCGTGC 2280
CTCTCTGCTGCCTTGGTGCCCCTGTGTCTGCCTCCTCTCC 2320
CTGCTTTTCAGCCTCTGTCTCCTTCCATCCCTGTCCCTGT 2360
CTGGCCTTAACTCTTCCCTCTCTTTCTTTTCTTTCTTTCT 2400
        2410        2420        2430        2440
TTCTTTTTTTTTAAGATAGAGTCTCACTCTGCTGCCCGGG 2440
CTAGAGTGCAGTGGTGGGATCTCGGCTCACTGCAACCTCT 2480
GCCTCCTGGGGTTCAAGTGATCCTCCCACCTCAGCCTCCT 2520
GAGTAGCTGGGATTACAGGCACCCGCCACCACGTCCAGCT 2560
AATTTTTATATTTTTAGTAGAGACGGGGTTTCACCATGTT 2600
```

Fig. 44B

hsFATP1 full lenght.DNA

```
        2610        2620        2630        2640
GGTCAGGCTGGTCTTGAACTCCTGACCTCAGGTGATCCGC 2640
TGGCCTCGGCCTCCCAGAGTGCTGGGATTATAGGCGTGAG 2680
CCTCTGGCCCGGCCTTTCCTTTTTCCTCTCCTCTCCTGCC 2720
GAGAGTGGAACACACGTGTCCTGGGAGCTGCATCTTGTGT 2760
AGGGTCCAGCTGCTTTTGGGGACTGCAGGAATCATCTCCC 2800
        2810        2820        2830        2840
CTGGGCCCTGGACTCGGACTGGGGCCTCCCCACCTCCCTC 2840
TCGGCTGTGCCTTACGGAGCCCCAATCCAGGCCTCCTGTG 2880
GCTGTTGGGTTCCAGATGCTGCAGCTCCATGTGACTTCCA 2920
AGCAGGCCCTCCGCCCTCCCTGCTGAATGGAGGAGCCGGG 2960
GGTCCCCCAGGCCAACTGGAAAATCTCCCAGGCTAGGCCA 3000
        3010        3020        3030        3040
ATTGCCTTTTGCACTTCCCCGTTCCTGTCACATTTCCCCA 3040
GCCCCACCTTCCCCTCCTGATGCCCTGAAAGCTTCCGGAA 3080
TTGACTGTGACCACTTGGATGTCACCACTGTCAGCCCCTG 3120
CCTTGATGTCCCCATTTAGCCATCTCCATGGAGCTCCTGC 3160
TGGAGGGCCCTGAACCCTGCACTGCGTGGCTGCCCAGCCA 3200
        3210        3220        3230        3240
GCTGCCTCCTGTCCTGGGAGGAGGCCTCCTGGGTGTCCTC 3240
ATCTGGTGTGTCTACTGGAGGGTCCCACAGGAGAGGCAGC 3280
AGAGGGGTCAGGGGAGGTCTCCTGCCGGGGGTTGGCCTCT 3320
CAAGCCTCAGGGGTTCTAGCCTGTTGAATATACCCCACCT 3360
GGTGGGTGGCCCCTCCGATGTCCCCACTGATGGCTCTGAC 3400
        3410        3420        3430        3440
ACCGTGTTGGTGGCGATGTCCCAGACAATCCCACCAGGAC 3440
GGCCCAGACATCCCTACTGGCTTCGCTGGTGGCTCATCTC 3480
GAACATCCACGCCAGCCTTTCTGGGGCCGGCCACCCAGGC 3520
CGCCTGTCCGTCTGTCCTCCCTCCAGCAGCACCCCCTGGC 3560
CCCTGGAGTGGTGGGGCCATGGCAAGAGACACCGTGGCGT 3600
        3610        3620        3630        3640
CTCATGTGAACTTTCCTGGGCACTGTGGTTTTATTTCCTA 3640
ATTGATTTAAGAAATAAACCTGAAGACCGTCTGGTGAAAA 3680
AAAAAAAAAAAAAA 3694
```

Fig. 44C

hsFATP1 full lenght.protein

```
         10         20         30         40
         |          |          |          |
MRAPGAGAASVVSLALLWLLGLPWTWSAAAALGVYVGSGG 40
WRFLRIVCKTARRDLFGLSVLIRVRLELRRHQRAGHTIPR 80
IFQAVVQRQPERLALVDAGTGECWTFAQLDAYSNAVANLF 120
RQLGFAPGDVVAIFLEGRPEFVGLWLGLAKAGMEAALLNV 160
NLRREPLAFCLGTSGAKALIFGGEMVAAVAEVSGHLGKSL 200
        210        220        230        240
         |          |          |          |
IKFCSGDLGPEGILPDTHLLDPLLKEASTAPLAQIPSKGM 240
DDRLFYIYTSGTTGLPKAAIVVHSRYYRMAAFGHHAYRMQ 280
AADVLYDCLPLYHSAGNIIGVGQCLIYGLTVVLRKKFSAS 320
RFWDDCIKYNCTVVQYIGEICRYLLKQPVREAERRHRVRL 360
AVGNGLRPAIWEEFTERFGVRQIGEFYGATECNCSIANMD 400
        410        420        430        440
         |          |          |          |
GKVGSCGFNSRILPHVYPIRLVKVNEDTMELLRDAQGLCI 440
PCQAGEPGLLVGQINQQDPLRRFDGYVSESATSKKIAHSV 480
FSKGDSAYLSGDVLVMDELGYMYFRDRSGDTFRWRGENVS 520
TTEVEGVLSRLLGQTDVAVYGVAVPGVEGKAGMAAVADPH 560
SLLDPNAIYQELQKVLAPYARPIFLRLLPQVDTTGTFKIQ 600
        610        620        630        640
         |          |          |          |
KTRLQREGFDPRQTSDRLFFLDLKQGHYLPLNEAVYTRIC 640
SGAFAL. 647
```

Fig. 45

```
        10         20         30         40
GGAATTCCAAAAAAAAAAAAATACGACTACACCTGCTCCGG 40
AGCCCGCGGCGGTACCTGCAGCGGAGGAGCTCTGTCTTCC 80
CCTTCATCTCACGCGAGCCCGGCGTCCCGCCGCGTGCGCC 120
CCGGCGCAGCCCGCCAGTCCGCCCGGAGCCCGCCCAGTCG 160
CCGCGCTGCACGCCCGGGGTGAACCCTCTGCCCTCGCTGG 200
        210        220        230        240
GACAGAGGGCCCCGCAGCCGTCATGCTTTCCGCCATCTAC 240
ACAGTCCTGGCGGGACTGCTGTTCCTGCCGCTCCTGGTGA 280
ACCTCTGCTGCCCATACTTCTTCCAGGACATAGGCTACTT 320
CTTGAAGGTGGCCGCCGTGGGCCGGAGGGTGCGCAGCTAC 360
GGGCAGCGGCGGCCGGCGCGCACCATCCTGCGGGCGTTCC 400
        410        420        430        440
TGGAGAAAGCGCGCCAGACGCCACACAAGCCTTTTCTGCT 440
CTTCCGCGACGAGACTCTCACCTACGCGCAGGTGGACCGG 480
CGCAGCAATCAAGTGGCCCGGGCGCTGCACGACCACCTCG 520
GCCTGCGCCAGGGAGACTGCGTGGCGCTCCTTATGGGTAA 560
CGAGCCGGCCTACGTGTGGCTGTGGCTGGGGCTGGTGAAG 600
        610        620        630        640
CTGGGCTGTGCCATGGCGTGCCTCAATTACAACATCCGCG 640
CGAAGTCCCTGCTGCACTGCTTCCAGTGCTGCGGGGCGAA 680
GGTGCTGCTGGTGTCGCCAGAACTACAAGCAGCTGTCGAA 720
GAGATACTGCCAAGCCTTAAAAAAGATGATGTGTCCATCT 760
ATTATGTGAGCAGAACTTCTAACACAGATGGGATTGACTC 800
        810        820        830        840
TTTCCTGGACAAAGTGGATGAAGTATCAACTGAACCTATC 840
CCAGAGTCATGGAGGTCTGAAGTCACTTTTTCCACTCCTG 880
CCTTATACATTTATACTTCTGGAACCACAGGTCTTCCAAA 920
AGCAGCCATGATCACTCATCAGCGCATATGGTATGGAACT 960
GGCCTCACTTTTGTAAGCGGATTGAAGGCAGATGATGTCA 1000
        1010       1020       1030       1040
TCTATATCACTCTGCCCTTTTACCACAGTGCTGCACTACT 1040
GATTGGCATTCACGGATGTATTGTGGCTGGTGCTACTCTT 1080
GCCTTGCGGACTAAATTTTCAGCCAGCCAGTTTTGGGATG 1120
ACTGCAGAAAATACAACGTCACTGTCATTCAGTATATCGG 1160
TGAACTGCTTCGGTATTTATGCAACTCACCACAGAAACCA 1200
```

Fig. 46 A

hsVLACS full lenght.DNA

```
        1210        1220        1230        1240
AATGACCGTGATCATAAAGTGAGACTGGCACTGGGAAATG 1240
GCTTACGAGGAGATGTGTGGAGACAATTTGTCAAGAGATT 1280
TGGGGACATATGCATCTATGAGTTCTATGCTGCCACTGAA 1320
GGCAATATTGGATTTATGAATTATGCGAGAAAAGTTGGTG 1360
CTGTTGGAAGAGTAAACTACCTACAGAAAAAAATCATAAC 1400
        1410        1420        1430        1440
TTATGACCTGATTAAATATGATGTGGAGAAAGATGAACCT 1440
GTCCGAGATGAAAATGGATATTGCGTCAGAGTTCCCAAAG 1480
GTGAAGTTGGACTTCTGGTTTGCAAAATCACACAACTTAC 1520
ACCATTTAATGGCTATGCTGGAGCAAAGGCTCAGACAGAG 1560
AAGAAAAAACTGAGAGATGTCTTTAAGAAAGGAGACCTCT 1600
        1610        1620        1630        1640
ATTTCAACAGTGGAGATCTCTTAATGGTTGACCATGAAAA 1640
TTTCATCTATTTCCACGACAGAGTTGGAGATACATTCCGG 1680
TGGAAAGGGGAAAATGTGGCCACCACTGAAGTTGCTGATA 1720
CAGTTGGACTGGTTGATTTTGTCCAAGAAGTAAATGTTTA 1760
TGGAGTGCATGTGCCAGATCATGAGGGTCGCATTGGCATG 1800
        1810        1820        1830        1840
GCCTCCATCAAAATGAAAGAAAACCATGAATTTGATGGAA 1840
AGAAACTCTTTCAGCACATTGCTGATTACCTACCTAGTTA 1880
TGCAAGGCCCCGGTTTCTAAGAATACAGGACACCATTGAG 1920
ATCACTGGAACTTTTAAACACCGCAAAATGACCCTGGTGG 1960
AGGAGGGCTTTAACCCTGCTGTCATCAAAGATGCCTTGTA 2000
        2010        2020        2030        2040
TTTCTTGGATGACACAGCAAAAATGTATGTGCCTATGACT 2040
GAGGACATCTATAATGCCATAAGTGCTAAAACCCTGAAAC 2080
TCTGAATATTCCCAGGAGGATAACTCAACATTTCCAGAAA 2120
GAAACTGAATGGACAGCCACTTGATATAATCCAACTTTAA 2160
TTTGATTGAAGATTGTGAGGAAATTTTGTAGGAAATTTGC 2200
        2210        2220        2230        2240
ATACCCGTAAAGGGAGACTTTTTTAAATAACAGTTGAGTC 2240
TTTGCAAGTAAAAAGATTTAGAGATTATTATTTTTCAGTG 2280
TGCACCTACTGTTTGTATTTGCAAACTGAGCTTGTTGGAG 2320
GGAAGGCATTATTTTTTAAAATACTTAGTAAATTAAATGA 2360
AC 2362
```

Fig. 46B

hsVLACS full lenght.protein

```
        10        20        30        40
MLSAIYTVLAGLLFLPLLVNLCCPYFFQDIGYFLKVAAVG 40
RRVRSYGQRRPARTILRAFLEKARQTPHKPFLLFRDETLT 80
YAQVDRRSNQVARALHDHLGLRQGDCVALLMGNEPAYVWL 120
WLGLVKLGCAMACLNYNIRAKSLLHCFQCCGAKVLLVSPE 160
LQAAVEEILPSLKKDDVSIYYVSRTSNTDGIDSFLDKVDE 200
       210       220       230       240
VSTEPIPESWRSEVTFSTPALYIYTSGTTGLPKAAMITHQ 240
RIWYGTGLTFVSGLKADDVIYITLPFYHSAALLIGIHGCI 280
VAGATLALRTKFSASQFWDDCRKYNVTVIQYIGELLRYLC 320
NSPQKPNDRDHKVRLALGNGLRGDVWRQFVKRFGDICIYE 360
FYAATEGNIGFMNYARKVGAVGRVNYLQKKIITYDLIKYD 400
       410       420       430       440
VEKDEPVRDENGYCVRVPKGEVGLLVCKITQLTPFNGYAG 440
AKAQTEKKKLRDVFKKGDLYFNSGDLLMVDHENFIYFHDR 480
VGDTFRWKGENVATTEVADTVGLVDFVQEVNVYGVHVPDH 520
EGRIGMASIKMKENHEFDGKKLFQHIADYLPSYARPRFLR 560
IQDTIEITGTFKHRKMTLVEEGFNPAVIKDALYFLDDTAK 600
       610       620       630       640
MYVPMTEDIYNAISAKTLKL. 621
```

Fig. 47

```
        10        20        30        40
AAGTTCTCGGCTGGTCAGTTCTGGGAAGATTGCCAGCAGC 40
ACAGGGTGACGGTGTTCCAGTACATTGGGGAGCTGTGCCG 80
ATACCTTGTCAACCAGCCCCCGAGCAAGGCAGAACGTGGC 120
CATAAGGTCCGGCTGGCAGTGGGCAGCGGGCTGCGCCCAG 160
ATACCTGGGAGCGTTTTGTGCGGCGCTTCGGGCCCCTGCA 200
       210       220       230       240
GGTGCTGGAGACATATGGACTGACAGAGGGCAACGTGGCC 240
ACCATCAACTACACAGGACAGCGGGGCGCTGTGGGGCGTG 280
CTTCCTGGCTTTACAAGCATATCTTCCCCTTCTCCTTGAT 320
TCGCTATGATGTCACCACAGGAGAGCCAATTCGGGACCCC 360
CAGGGGCACTGTATGGCCACATCTCCAGGTGAGCCAGGGC 400
       410       420       430       440
TGCTGGTGGCCCCGGTAAGCCAGCAGTCCCCATTCCTGGG 440
CTATGCTGGCGGGCCAGAGCTGGCCCAGGGGAAGTTGCTA 480
AAGGATGTCTTCCGGCCTGGGGATGTTTTCTTCAACACTG 520
GGGACCTGCTGGTCTGCGATGACCAAGGTTTTCTCCGCTT 560
CCATGATCGTACTGGAGACACCTTCAGGTGGAAGGGGGAG 600
       610       620       630       640
AATGTGGCCACAACCGAGGTGGCAGAGGTCTTCGAGGCCC 640
TAGATTTTCTTCAGGAGGTGAACGTCTATGGAGTCACTGT 680
GCCAGGGCATGAAGGCAGGGCTGGAATGGCAGCCCTAGTT 720
CTGCGTCCCCCCCACGCTTTGGACCTTATGCAGCTCTACA 760
CCCACGTGTCTGAGAACTTGCCACCTTATGCCCGGCCCCG 800
       810       820       830       840
ATTCCTCAGGCTCCAGGAGTCTTTGGCCACCACAGAGACC 840
TTCAAACAGCAGAAAGTTCGGATGGCAAATGAGGGCTTCG 880
ACCCCAGCACCCTGTCTGACCCACTGTACGTTCTGGACCA 920
GGCTGTAGGTGCCTACCTGCCCCTCACAACTGCCCGGTAC 960
AGCGCCCTCCTGGCAGGAAACCTTCGAATCTGAGAACTTC 1000
      1010      1020      1030      1040
CACACCTGAGGCACCTGAGAGAGGAACTCTGTGGGGTGGG 1040
GGCCGTTGCAGGTGTACTGGGCTGTCAGGGATCTTTTCTA 1080
TACCAGAACTGCGGTCACTATTTTGTAATAAATGTGGCTG 1120
GAGCTGATCCAGCTGTCTCTGACAAAAAAAAAAAAAAAAA 1160
AAAGGGCGGCCGC 1173
```

Fig. 48

hsFATP3partial.protein

```
          10        20        30        40
KFSAGQFWEDCQQHRVTVFQYIGELCRYLVNQPPSKAERG 40
HKVRLAVGSGLRPDTWERFVRRFGPLQVLETYGLTEGNVA 80
TINYTGQRGAVGRASWLYKHIFPFSLIRYDVTTGEPIRDP 120
QGHCMATSPGEPGLLVAPVSQQSPFLGYAGGPELAQGKLL 160
KDVFRPGDVFFNTGDLLVCDDQGFLRFHDRTGDTFRWKGE 200
         210       220       230       240
NVATTEVAEVFEALDFLQEVNVYGVTVPGHEGRAGMAALV 240
LRPPHALDLMQLYTHVSENLPPYARPRFLRLQESLATTET 280
FKQQKVRMANEGFDPSTLSDPLYVLDQAVGAYLPLTTARY 320
SALLAGNLRI. 331
```

Fig. 49

hsFATP4 full length

```
        10        20        30        40
CGACCCACGCGTCCGGGCGGGCGGGGCCGGGCGGCGGGCG 40
GGGCTGGCGGGGCGGCCGGGCCATGCAGGGCGCAGAGCCG 80
GCTAAACCCTGCTGAGACCCGGCTCCGTGCGTCCAGGGGC 120
GGCTAATGCCCCTCACGCTGTCTACGCTGCTGCAACCGGG 160
CCGCATCTGGACGGGGCGCCGCGCGGCGGAGCCGACGCCG 200
       210       220       230       240
GGCCACAATGCTGCTTGGAGCCTCTCTGGTGGGGGTGCTG 240
CTGTTCTCCAAGCTGGTGCTGAAACTGCCCTGGACCCAGG 280
TGGGATTCTCCCTGTTGTTCCTCTACTTGGGATCTGGCGG 320
CTGGCGCTTCATCCGGGTCTTCATCAAGACCATCAGGCGC 360
GATATCTTTGGCGGCCTGGTCCTCCTGAAGGTGAAGGCAA 400
       410       420       430       440
AGGTGCGACAGTGCCTGCAGGAGCGGCGGACAGTGCCCAT 440
TTTGTTTGCCTCTACCGTTCGGCGCCACCCCGACAAGACG 480
GCCCTGATCTTCGAGGGCACAGATACCCACTGGACCTTCC 520
GCCAGCTGGATGAGTACTCAAGCAGTGTAGCCAACTTCCT 560
GCAGGCCCGGGGCCTGGCCTCGGGCGATGTGGCTGCCATC 600
       610       620       630       640
TTCATGGAGAACCGCAATGAGTTCGTGGGCCTATGGCTGG 640
GCATGGCCAAGCTCGGTGTGGAGGCAGCCCTCATCAACAC 680
CAACCTGCGGCGGGATGCTCTGCTCCACTGCCTCACCACC 720
TCGCGCGCACGGGCCCTTGTCTTTGGCAGCGAAATGGCCT 760
CAGCCATCTGTGAGGTCCATGCCAGCCTGGACCCCTCGCT 800
       810       820       830       840
CAGCCTCTTCTGCTCTGGCTCCTGGGAGCCCGGTGCGGTG 840
CCTCCAAGCACAGAACACCTGGACCCTCTGCTGAAAGATG 880
CTCCCAAGCACCTTCCCAGTTGCCCTGACAAGGGCTTCAC 920
AGATAAACTGTTCTACATCTACACATCCGGCACCACAGGG 960
CTGCCCAAGGCCGCCATCGTGGTGCACAGCAGGTATTACC 1000
      1010      1020      1030      1040
GCATGGCTGCCCTGGTGTACTATGGATTCCGCATGCGGCC 1040
CAACGACATCGTCTATGACTGCCTCCCCCTCTACCACTCA 1080
GCAGGAAACATCGTGGGAATCGGCCAGTGCCTGCTGCATG 1120
GCATGACGGTGGTGATTCGGAAGAAGTTCTCAGCCTCCCG 1160
GTTCTGGGACGATTGTATCAAGTACAACTGCACGATTGTG 1200
```

Fig. 50A

hsFATP4 full length

```
        1210        1220        1230        1240
CAGTACATTGGTGAACTGTGCCGCTACCTCCTGAACCAGC 1240
CACCGCGGGAGGCAGAAAACCAGCACCAGGTTCGCATGGC 1280
ACTAGGCAATGGCCTCCGGCAGTCCATCTGGACCAACTTT 1320
TCCAGCCGCTTCCACATACCCCAGGTGGCTGAGTTCTACG 1360
GGGCCACAGAGTGCAACTGTAGCCTGGGCAACTTCGACAG 1400
        1410        1420        1430        1440
CCAGGTGGGGGCCTGTGGTTTCAATAGCCGCATCCTGTCC 1440
TTCGTGTACCCCATCCGGTTGGTACGTGTCAACGAGGACA 1480
CCATGGAGCTGATCCGGGGGCCCGACGGCGTCTGCATTCC 1520
CTGCCAGCCAGGTGAGCCGGGCCAGCTGGTGGGCCGCATC 1560
ATCCAGAAAGACCCCCTGCGCCGCTTCGATGGCTACCTCA 1600
        1610        1620        1630        1640
ACCAGGGCGCCAACAACAAGAAGATTGCCAAGGATGTCTT 1640
CAAGAAGGGGGACCAGGCCTACCTTACTGGTGATGTGCTG 1680
GTGATGGACGAGCTGGGCTACCTGTACTTCCGAGACCGCA 1720
CTGGGGACACGTTCCGCTGGAAAGGTGAGAACGTGTCCAC 1760
CACCGAGGTGGAAGGCACACTCAGCCGCCTGCTGGACATG 1800
        1810        1820        1830        1840
GCTGACGTGGCCGTGTATGGTGTCGAGGTGCCAGGAACCG 1840
AGGGCCGGGCCGGAATGGCTGCTGTGGCCAGCCCCACTGG 1880
CAACTGTGACCTGGAGCGCTTTGCTCAGGTCTTGGAGAAG 1920
GAACTGCCCCTGTATGCGCGCCCCATCTTCCTGCGCCTCC 1960
TGCCTGAGCTGCACAAAACAGGAACCTACAAGTTCCAGAA 2000
        2010        2020        2030        2040
GACAGAGCTACGGAAGGAGGGCTTTGACCCGGCTATTGTG 2040
AAAGACCCGCTGTTCTATCTAGATGCCCAGAAGGGCCGCT 2080
ACGTCCCGCTGGACCAAGAGGCCTACAGCCGCATCCAGGC 2120
AGGCGAGGAGAAGCTGTGATTCCCCCCATCCCTCTGAGGG 2160
CCGGCGGATGCTGGATCCGGAGCCCCAGGTTCCGCCCCAG 2200
        2210        2220        2230        2240
AGCGGTCCTGGACAAGGCCAGACCAAAGCAAGCAGGGCCT 2240
GGCACCTCCATCCTGAGGTGCTGCCCCTCCATCCAAAACT 2280
GCCAAGTGACTCATTGCCTTCCCAACCCTTCCAGAGGCTT 2320
TCTGTGAAAGTCTCATGTCCAAGTTCCGTCTTCTGGGCTG 2360
GGCAGGCCCTCTGGTTCCCAGGCTGAGACTGACGGGTTTT 2400
        2410        2420        2430        2440
CTCAGGATGATGTCTTGGGTGAGGGTAGGGAGAGGACAAG 2440
GGGTCACCGAGCCCTTCCCAGAGAGCAGGGAGCTTATAAA 2480
TGGAACCAGAGCAGAAGTCCCCAGACTCAGGAAGTCAACA 2520
GAGTGGGCAGGGACAGTGGTAGCATCCATCTGGTGGCCAA 2560
AGAGAATCGTAGCCCCAGAGCTGCCCAAGTTCACTGGGCT 2600
```

Fig. 50B

hsFATP4 full length

```
         2610      2620      2630      2640
CCACCCCCACCTCCAGGAGGGGAGGAGAGGACCTGACATC 2640
TGTAGGTGGCCCCTGATGCCCCATCTACAGCAGGAGGTCA 2680
GGACCACGCCCCTGGCCTCTCCCCACTCCCCCATCCTCCT 2720
CCCTGGGTGGCTGCCTGATTATCCCTCAGGCAGGGCCTCT 2760
CAGTCCTTGTGGGTCTGTGTCACCTCCATCTCAGTCTTGG 2800
         2810      2820      2830      2840
CCTGGCTATGAGGGGAGGAGGAATGGGAGAGGGGGCTCAG 2840
GGGCCAATAAACTCTGCCTTGAGTCCTCCTAAAAAAAAAA 2880
AAAAAAAAAAAAAAAAAAAAAAAAAAA 2907
```

Fig. 50C

hsFATP4 full length. protein

```
         10        20        30        40
         .....|.....|.....|.....|.....|.....|.....|.....|
MLLGASLVGVLLFSKLVLKLPWTQVGFSLLFLYLGSGGWR 40
FIRVFIKTIRRDIFGGLVLLKVKAKVRQCLQERRTVPILF 80
ASTVRRHPDKTALIFEGTDTHWTFRQLDEYSSSVANFLQA 120
RGLASGDVAAIFMENRNEFVGLWLGMAKLGVEAALINTNL 160
RRDALLHCLTTSRARALVFGSEMASAICEVHASLDPSLSL 200
         210       220       230       240
         .....|.....|.....|.....|.....|.....|.....|.....|
FCSGSWEPGAVPPSTEHLDPLLKDAPKHLPSCPDKGFTDK 240
LFYIYTSGTTGLPKAAIVVHSRYYRMAALVYYGFRMRPND 280
IVYDCLPLYHSAGNIVGIGQCLLHGMTVVIRKKFSASRFW 320
DDCIKYNCTIVQYIGELCRYLLNQPPREAENQHQVRMALG 360
NGLRQSIWTNFSSRFHIPQVAEFYGATECNCSLGNFDSQV 400
         410       420       430       440
         .....|.....|.....|.....|.....|.....|.....|.....|
GACGFNSRILSFVYPIRLVRVNEDTMELIRGPDGVCIPCQ 440
PGEPGQLVGRIIQKDPLRRFDGYLNQGANNKKIAKDVFKK 480
GDQAYLTGDVLVMDELGYLYFRDRTGDTFRWKGENVSTTE 520
VEGTLSRLLDMADVAVYGVEVPGTEGRAGMAAVASPTGNC 560
DLERFAQVLEKELPLYARPIFLRLLPELHKTGTYKFQKTE 600
         610       620       630       640
         .....|.....|.....|.....|.....|.....|.....|.....|
LRKEGFDPAIVKDPLFYLDAQKGRYVPLDQEAYSRIQAGE 640
EKL 643
```

Fig. 51

>hsFATP5(partial)
GTCGTTGGGATCCTCGGCTGCTTAGATCTCGGAGCCACCTGTGTTCTGGCCCCCAAG
TTCTCTACTTCCTGCTTCTGGGA
TGACTGTCGGCAGCATGGCGTGACAGTGATCCTGTATGTGGGCGAGCTCCTGCGATA
CTTGTGTAACATTCCCCAGCAAC
CAGAGGACCGGACACATACAGTCCGCCTGGCAATGGGCAATGGACTACGGGCTGAT
GTGTGGGGAGACCTTCCAGCAGCG
TTTCGGTCCTATTTCGGATCTNGGGAAGTCTTACGGGCTTCCACAGAAGGGCAACAT
GGGGCTTTAGTTCAAATATTGTT
GGGGGCGCTGCGGGGCCCTGGGGGCAAAGATGGAGCTTGCCTCCTCCGAATGCTGT
CCCCCTTTGAGCTGGTGCAGTTCG
ACATGGAGGCGGCGGAGCCTGTGAGGGACAATCAGGGCTTCTGCATCCCTGTAGGG
CTAGGGGAGCCGGGGCTGCTGTTG
ACCAAGGTGGTAAGCCAGCAACCCTTCGTGGGCTACCGCGGCCCCCGAGAGCTGTC
GGAACGGAAGCTGGTGCGCAACGT
GCGGCAATCGGGCGACGTTTACTACAACACCGGGGACGTACTGGCCATGGACCGCG
AAGGCTTCCTCTACTTCCGCGACC
GACTCGGGGACACCTTCCGATGGAAGGGCGAGAACGTGTCCACGCACGAGGTGGAG
GGCGTGTTGTCGCAGGTGGACTTC
TTGCAACAGGTTAACGTGTATGGCGTGTGCGTGCCAGGTTGTGAGGGTAAGGTGGGC
ATGGCTGCTGTGGCATTAGCCCC
CGGCCAGACTTTCGACGGGGAGAAGTTGTACCAGCACGTTCGCGCTTGGCTCCCTGC
CTACGCTACCCCCCATTTCATCC
GCATCCAGGACGCCATGGAGGTCACCAGCACGTTCAAACTGATGAAGACCCGGTTG
GTGCGTGAGGGCTTCAATGTGGGG
ATCGTGGTTGACCCTCTGTTTGTACTGGACAACCGGGCCCAGTCCTTCCGGCCCCTG
ACGGCAGAAATGTACCAGGCTGT
GTGTGAGGGAACCTGGAGGCTCTGATCACCTGGCCAACCCACTGGGGTAGGGATCA
AAGCCAGCCACCCCCACCCCAACA
CACTCGGTGTCCCTTTCATCCTGGGCCTGTGTGAATCCCAGCCTGGCCATACCCTCA
ACCTCAGTGGGCTGGAAATGACA
GTGGGCCCTGTAGCAGTGGCAGAATAAACTCAGMTGYGTTCACAGAAA

Fig. 52

hsFATP5partial.protein

...

```
         10        20        30        40
VVGILGCLDLGATCVLAPKFSTSCFWDDCRQHGVTVILYV 40
GELLRYLCNIPQQPEDRTHTVRLAMGNGLRADVWGDLPAA 80
FRSYFGSXEVLRASTEGQHGALVQILLGALRGPGGKDGAC 120
LLRMLSPFELVQFDMEAAEPVRDNQGFCIPVGLGEPGLLL 160
TKVVSQQPFVGYRGPRELSERKLVRNVRQSGDVYYNTGDV 200
        210       220       230       240
LAMDREGFLYFRDRLGDTFRWKGENVSTHEVEGVLSQVDF 240
LQQVNVYGVCVPGCEGKVGMAAVALAPGQTFDGEKLYQHV 280
RAWLPAYATPHFIRIQDAMEVTSTFKLMKTRLVREGFNVG 320
IVVDPLFVLDNRAQSFRPLTAEMYQAVCEGTWRL 354
```

Fig. 53

```
     1         10         20         30         40
     |         |          |          |          |
AACGGCAAGTAAGCGCAACGCAATTAATGTGAGTAGCTCA 40
CTCATTAGGCACCCCAGGCTTTACACTTTATGCTTCCGGG 80
CTCGTATGTTGTGTGGAATTGTGAGCGGATACCAATTTCA 120
CACAGGAACCAGCTATGACATGATTACGAATTTAATACGA 160
CTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGG 200
   210        220        230        240
   |          |          |          |
CACGAGGGGTGCTGAGCCCCTGCGCGGTTTCTGGTGCGTA 240
GAGACTGTAAATCGCTGCGCTTCTCAGTCATCATCATCCC 280
AGCTTTTCCCGGCTCGAATTCAGCCTCCAACTCAAGCTCG 320
CGGGAAAGACTACCTGAGAGGAGAAAAGCTTCTGTCCCTG 360
GACCTTCTTCTGAGGGTGGAGTCGGAGGCTCCCTGCTTTC 400
   410        420        430        440
   |          |          |          |
CAGCCGCCCAGTGACCCAAGCTTAATCTTCAGCACCACTT 440
GGGGCGACCTTTTCGGTGCAAACCTACGATTCTGTTTCTC 480
AGGATTCCTCCCCATCCCGCTTCGCCCCGGAAAAGCTGAC 520
AAGAACTTCAGGTGTAAGCCCTGAGTAGTGAGGATCTGCG 560
GTCTCCGTGGAGAGCTGTGCCTGGAAGAGAAGGACGCTGG 600
   610        620        630        640
   |          |          |          |
TGGGGGCTGAGATCAGAGCTGTCTTCTGGCCCAGTTGCCC 640
CCATGCTTCTGTCATGGCTAACAGTTCTAGGGGCTGGAAT 680
GGTCGTCCTGCACTTCTTGCAGAAACTCCTGTTCCCTTAC 720
TTTTGGGATGACTTCTGGTTCGTGTTGAAGGTGGTGCTCA 760
TTATAATTCGGCTGAAGAAGTATGAAAAGAGAGGGGAGCT 800
   810        820        830        840
   |          |          |          |
GGTGACTGTGCTGGATAAATTCTTGAGTCATGCCAAAAGA 840
CAACCTCGGAAACCTTTCATCATCTATGAGGGAGACATCT 880
ACACCTATCAGGATGTAGACAAAAGGAGCAGCAGAGTGGC 920
CCATGTCTTCCTGAACCATTCCTCTCTGAAAAAGGGGGAC 960
ACGGTGGCTCTGCTGATGAGCAATGAGCCGGACTTCGTTC 1000
  1010       1020       1030       1040
   |          |          |          |
ACGTGTGGTTCGGCCTCGCCAAGCTGGGCTGCGTGGTGGC 1040
CTTTCTCAACACCAACATTCGCTCCAACTCCCTCCTGAAT 1080
TGCATCCGCGCCTGTGGGCCCAGAGCCCTAGTGGTGGGCG 1120
CAGATTTGCTTGGAACGGTAGAAGAAATCCTTCCAAGCCT 1160
CTCAGAAAATATCAGTGTTTGGGGGATGAAAGATTCTGTT 1200
```

Fig. 54A

```
        1210      1220      1230      1240
CCACAAGGTGTAATTTCACTCAAAGAAAAACTGAGCACCT 1240
CACCTGATGAGCCCGTGCCACGCAGCCACCATGTTGTCTC 1280
ACTCCTCAAGTCTACTTGTCTTTACATTTTTACCTCTGGA 1320
ACAACAGGTCTACCAAAAGCAGCTGTGATTAGTCAGCTGC 1360
AGGTTTTAAGGGGTTCTGCTGTCCTGTGGGCTTTTGGTTG 1400
        1410      1420      1430      1440
TACTGCTCATGACATTGTTTATATAACCCTTCCTCTGTAT 1440
CATAGTTCAGCAGCTATCCTGGGAATTTCTGGATGTGTTG 1480
AGTTGGGTGCCACTTGTGTGTTAAAGAAGAAATTTTCAGC 1520
AAGCCAGTTTTGGAGTGACTGCAAGAAGTATGATGTGACT 1560
GTGTTTCAGTATATTGGAGAACTTTGTCGCTACCTTTGCA 1600
        1610      1620      1630      1640
AACAATCTAAGAGAGAAGGAGAAAAGGATCATAAGGTGCG 1640
TTTGGCAATTGGAAATGGCATACGGAGTGATGTATGGAGA 1680
GAATTTTTAGACAGATTTGGAAATATAAAGGTGTGTGAAC 1720
TTTATGCAGCTACCGAATCAAGCATATCTTTCATGAACTA 1760
CACTGGGAGAATTGGAGCAATTGGGAGAACAAATTTGTTT 1800
        1810      1820      1830      1840
TACAAACTTCTTTCCACTTTTGACTTAATAAAGTATGACT 1840
TTCAGAAAGATGAACCCATGAGAAATGAGCAGGGTTGGTG 1880
TATTCATGTGAAAAAAGGAGAACCTGGACTTCTCATTTCT 1920
CGAGTGAATGCAAAAAATCCCTTCTTTGGCTATGCTGGGC 1960
CTTATAAGCACACAAAAGACAAATTGCTTTGTGATGTTTT 2000
        2010      2020      2030      2040
TAAGAAGGGAGATGTTTACCTTAATACTGGAGACTTAATA 2040
GTCCAGGATCAGGACAATTTCCTTTATTTTTGGGACCGTA 2080
CTGGAGACACTTTCAGATGGAAAGGAGAAAATGTCGCAAC 2120
CACTGAGGTTGCTGATGTTATTGGAATGTTGGATTTCATA 2160
CAGGAAGCAAACGTCTATGGTGTGGCTATATCAGGTTATG 2200
        2210      2220      2230      2240
AAGGAAGAGCAGGAATGGCTTCTATTATTTTAAAACCAAA 2240
TACATCTTTAGATTTGGAAAAAGTTTATGAACAAGTTGTA 2280
ACATTTCTACCAGCTTATGCTTGTCCACGATTTTTAAGAA 2320
TTCAGGAAAAAAATGGAAGCAACAGGAACATTCAAACTATT 2360
GAAGCATCAGTTGGTGGAAGATGGATTTAATCCACTGAAA 2400
        2410      2420      2430      2440
ATTTCTGAACCACTTTACTTCATGGATAACTTGAAAAAGT 2440
CTTATGTTCTACTGACCAGGGAACTTTATGATCAAATAAT 2480
GTTAGGGGAAATAAAACTTTAAGATTTTTATATCTAGAAC 2520
TTTCATATGCTTTCTTAGGAAGAGTGAGAGGGGGGTATAT 2560
GATTCTTTATGAAATGGGGAAAGGGAGCTAACATTAATTA 2600
```

Fig. 54B

```
        2610        2620        2630        2640
TGCATGTACTATATTTCCTTAATATGAGAGATAATTTTTT 2640
AATTGCATAAGAATTTTAATTTCTTTTAATTGATATAAAC 2680
ATTAGTTGATTATTCTTTTTATCTATTTGGAGATTCAGTG 2720
CATAACTAAGTATTTTCCTTAATACTAAAGATTTTAAATA 2760
ATAAATAGTGGCTAGCGGTTTGGACAATCACTAAAAATGT 2800
        2810        2820        2830        2840
ACTTTCTAATAAGTAAAATTTCTAATTTTGAATAAAAGAT 2840
TAAATTTTACTGAAAAAAAAAAAAAAAAAAAAAATTGGCG 2880
GCCGC 2885
```

Fig. 54C

hsFATP6 full lenght.protein

```
         10        20        30        40
MLLSWLTVLGAGMVVLHFLQKLLFPYFWDDFWFVLKVVLI 40
IIRLKKYEKRGELVTVLDKFLSHAKRQPRKPFIIYEGDIY 80
TYQDVDKRSSRVAHVFLNHSSLKKGDTVALLMSNEPDFVH 120
VWFGLAKLGCVVAFLNTNIRSNSLLNCIRACGPRALVVGA 160
DLLGTVEEILPSLSENISVWGMKDSVPQGVISLKEKLSTS 200
        210       220       230       240
PDEPVPRSHHVVSLLKSTCLYIFTSGTTGLPKAAVISQLQ 240
VLRGSAVLWAFGCTAHDIVYITLPLYHSSAAILGISGCVE 280
LGATCVLKKKFSASQFWSDCKKYDVTVFQYIGELCRYLCK 320
QSKREGEKDHKVRLAIGNGIRSDVWREFLDRFGNIKVCEL 360
YAATESSISFMNYTGRIGAIGRTNLFYKLLSTFDLIKYDF 400
        410       420       430       440
QKDEPMRNEQGWCIHVKKGEPGLLISRVNAKNPFFGYAGP 440
YKHTKDKLLCDVFKKGDVYLNTGDLIVQDQDNFLYFWDRT 480
GDTFRWKGENVATTEVADVIGMLDFIQEANVYGVAISGYE 520
GRAGMASIILKPNTSLDLEKVYEQVVTFLPAYACPRFLRI 560
QEKMEATGTFKLLKHQLVEDGFNPLKISEPLYFMDNLKKS 600
        610       620       630       640
YVLLTRELYDQIMLGEIKL. 620
```

Fig. 55

```
         10         20         30         40
         |          |          |          |
AAGTTCCCACTCCAGACTTCTGCGAGAACCCGTGAGGAAG 40
CAGCGAGAACCGGGGGTTTGCAAGCCAGAGAAGGATGCGG 80
ACTCCGGGAGCAGGAACAGCCTCTGTGGCCTCATTGGGGC 120
TGCTTTGGCTTCTGGGACTTCCGTGGACCTGGAGCGCGGC 160
GGCGGCGTTCGGTGTGTACGTGGGTAGCGGTGGCTGGCGA 200
        210        220        230        240
        |          |          |          |
TTTCTGCGTATCGTCTGCAAGACGGCGAGGCGAGACCTCT 240
TTGGCCTCTCTGTTCTGATCCGCGTGCGGCTAGAGCTACG 280
ACGACACCGGCGAGCAGGAGACACGATCCCACGCATCTTC 320
CAGGCCGTGGCCCAGCGACAGCCGGAGCGCCTGGCGCTGG 360
TAGATGCGAGTAGCGGTATCTGCTGGACCTTCGCACAGCT 400
        410        420        430        440
        |          |          |          |
AGACACCTACTCCAATGCTGTGGCCAATCTGTTCCTCCAG 440
CTGGGCTTTGCGCCAGGCGATGTGGTGGCTGTGTTCCTGG 480
AAGGCCGGCCCGAGTTCGTGGGACTGTGGCTGGGCCTGGC 520
CAAGGCCGGTGTAGTGGCTGCGCTTCTCAATGTCAACCTG 560
AGGCGGGAGCCCCTTGCCTTCTGCTTGGGCACATCAGCTG 600
        610        620        630        640
        |          |          |          |
CCAAGGCCCTCATTTATGGCGGGGAGATGGCAGCGGCGGT 640
GGCGGAGGTGAGTGAGCAGCTGGGGAAGAGCCTGCTCAAG 680
TTCTGCTCTGGAGATCTGGGGCCTGAGAGCGTCCTGCCTG 720
ACACGCAGCTTCTGGACCCCATGCTTGCTGAGGCGCCCAC 760
CACACCCCTGGCACAGGCCCCAGGCAAGGGCATGGATGAT 800
        810        820        830        840
        |          |          |          |
CGGCTATTTTACATCTATACTTCTGGGACCACCGGACTTC 840
CTAAGGCGGCCATTGTGGTGCACAGCAGGTACTACCGCAT 880
CGCAGCCTTCGGCCACCATTCCTACAGCATGCGGGCCAAC 920
GATGTGCTCTATGACTGCCTACCTCTCTACCACTCAGCAG 960
GGAACATCATGGGCGTGGGACAGTGTATCATCTACGGGTT 1000
        1010       1020       1030       1040
        |          |          |          |
AACGGTGGTACTGCGCAAGAAGTTCTCCGCCAGCCGCTTC 1040
TGGGACGACTGTGTCAAATATAATTGCACGGTAGTGCAGT 1080
ACATCGGTGAAATATGCCGCTACCTGCTAAGGCAGCCGGT 1120
TCGCGATGTAGAGCGGCGGCACCGCGTGCGCCTGGCCGTG 1160
GGTAACGGACTGCGGCCAGCCATCTGGGAGGAGTTCACGC 1200
```

Fig. 56A

rnFATP1 full length.DNA

```
      1210      1220      1230      1240
AGGGTTTCGGTGTGCGACAGATTGGCGAGTTCTACGGCGC 1240
CACCGAATGCAACTGCAGCATTGCCAACATGGACGGCAAG 1280
GTCGGCTCCTGCGGCTTCAACAGCCGTATCCTCACGCATG 1320
TGTACCCCATCCGTCTGGTCAAGGTCAACGAGGACACGAT 1360
GGAGCCACTGAGGGACTCCCAAGGCCTCTGCATCCCGTGC 1400
      1410      1420      1430      1440
CAGCCCGGGGAACCTGGGCTTCTCGTGGGCCAGATCAACC 1440
AGCAAGACCCTCTGCGGCGCTTCGATGGCTATGTTAGTGA 1480
CAGCGCCACCAACAAGAAGATTGCCCACAGCGTGTTCCGA 1520
AAGGGGGACAGCGCCTACCTTTCAGGTGACGTGCTAGTGA 1560
TGGACGAGCTGGGGTACATGTACTTCCGTGACCGCAGCGG 1600
      1610      1620      1630      1640
GGATACCTTCCGATGGCGCGGCGAGAACGTATCCACCACG 1640
GAGGTGGAAGCCGTGCTGAGCCGCCTGTTGGGCCAGACGG 1680
ACGTGGCTGTGTATGGAGTGGCTGTGCCAGGAGTGGAGGG 1720
GAAAAGCGGCATGGCGGCCATTGCAGACCCCCACAACCAG 1760
CTGGACCCTAACTCAATGTACCAGGAATTGCAGAAGGTTC 1800
      1810      1820      1830      1840
TTGCATCCTATGCCCAGCCCATCTTCCTGCGTCTTCTGCC 1840
CCAAGTGGATACAACAGGCACCTTCAAGATCCAGAAGACC 1880
CGACTACAGCGTGAAGGCTTTGACCCCCGCCAGACCTCAG 1920
ACCGGCTCTTCTTTCTAGACCTGAAACAGGGACGCTACCT 1960
ACCCCTGGATGAGAGAGTCCATGCCCGCATCTGCGCAGGC 2000
      2010      2020      2030      2040
GACTTCTCACTCTGAGCCTGGTGAGTGGGATGGCCCTGGA 2040
CTTGTGAGACCAGGGAGCCGGACACCCCTGTTCAGGTGTT 2080
TCTCCTGCCTGGCCACGTGGCCAGCAGCACCTGTGGGTGC 2120
AGGAAACTGGAACCTGAGTGGCCGGGTGTCCCTTTCCTAC 2160
AACCCACCATGCACACATCTAGCCTCTGCCTTGGTCTTTT 2200
      2210      2220      2230      2240
TCTCCATCTCTTTCCTCCGTGCCCAGCAGGAGCCCCACAG 2240
ACACATTGGCTGCTGTGTCCTGCAGTGGGACCGGTGTCTA 2280
GGGGTCCATGCTGCAGGCTGTGACCCGCACTGGTGCCCAC 2320
CTCCCTTCCCCATTGTGCCTTAGGTTCCTCCACTGTGCGC 2360
CGGTGAAGCAAGTGGGGACCCACATAGCTGTTGTCCCTGC 2400
      2410      2420      2430      2440
TGAGGGTTGGTAGCAAATGCACCCTCATGTCAGCTGGGAG 2440
ACACATGCAGTCTCCCACTGACCCCCAATCAACTGAAGAT 2480
ACTGTTTTGTATTATTGTTTTGAGATAGGGTCTCACTGTG 2520
GAGGCCAAGCTGGCCTCAGGCTCACCACTCTACTGCCTCC 2560
GGGCACCAGCCTGCAGTTTGATGACATGTATGCACTATTG 2600
```

Fig. 56B

mFATP1 full length.DNA

2610 2620 2630 2640

```
TTCTAAGGGTCTTCTGAGTCCCTGCTTTCCCCTCATGTCC 2640
TAAAACCTTCCAGAACTGACTCTGATCACTTGGATGTAGC 2680
TAGTGTTGGCCCTGCCCACGTGTGTCAATTCAGGGGTCCC 2720
CAGGCATCATCTCTGGAGGCCCTAACCTTGGCAAAGCTTG 2760
GATGTCCTCACATCACAGCAGGAGACCCAGGAAGGTTGCT 2800
```

2810 2820 2830 2840

```
GTGGTGTCTCTTGGGCACCCCTGGCGGCAGCCGTGGACAT 2840
GCTTCCCTGCTGTGATAGCCCAAACTGTTGCCTATGACAT 2880
TTGAGGTCTACCCTTCTGGCTGCCATGGTCCCCATTGAGA 2920
TCTTTGGTGACTCACCTCAGCCACCAAGCCAGGCCTCTGC 2960
CTTCCTTCAGCTCTAAGGGCATGAAGGGTGTGGACAGAGC 3000
```

3010 3020 3030 3040

```
AGCCACAGGCTGCCCACAGTCACCCACATGCAAGTGTTAT 3040
TTCCTTGTTTGTTTTAAAAAAATAAACATGCTGAGCCTTG 3080
AAAAAAAAAAAAAAAAAA 3098
```

Fig. 56C

mFATP1 full lenght.protein

```
         10         20         30         40
MRTPGAGTASVASLGLLWLLGLPWTWSAAAAFGVYVGSGG 40
WRFLRIVCKTARRDLFGLSVLIRVRLELRRHRRAGDTIPR 80
IFQAVAQRQPERLALVDASSGICWTFAQLDTYSNAVANLF 120
LQLGFAPGDVVAVFLEGRPEFVGLWLGLAKAGVVAALLNV 160
NLRREPLAFCLGTSAAKALIYGGEMAAAVAEVSEQLGKSL 200
        210        220        230        240
LKFCSGDLGPESVLPDTQLLDPMLAEAPTTPLAQAPGKGM 240
DDRLFYIYTSGTTGLPKAAIVVHSRYYRIAAFGHHSYSMR 280
ANDVLYDCLPLYHSAGNIMGVGQCIIYGLTVVLRKKFSAS 320
RFWDDCVKYNCTVVQYIGEICRYLLRQPVRDVERRHRVRL 360
AVGNGLRPAIWEEFTQGFGVRQIGEFYGATECNCSIANMD 400
        410        420        430        440
GKVGSCGFNSRILTHVYPIRLVKVNEDTMEPLRDSQGLCI 440
PCQPGEPGLLVGQINQQDPLRRFDGYVSDSATNKKIAHSV 480
FRKGDSAYLSGDVLVMDELGYMYFRDRSGDTFRWRGENVS 520
TTEVEAVLSRLLGQTDVAVYGVAVPGVEGKSGMAAIADPH 560
NQLDPNSMYQELQKVLASYAQPIFLRLLPQVDTTGTFKIQ 600
        610        620        630        640
KTRLQREGFDPRQTSDRLFFLDLKQGRYLPLDERVHARIC 640
AGDFSL.  647
```

Fig. 57

```
         10         20         30         40
....|....|....|....|....|....|....|....|....|
GACACAGTACTGCCGATGTTGGACAGAGGATCGCTTAACA 40
GAACGAAATCTCAAAACAAATTAACAGGACCCGGTTGCTT 80
GATTTCCCAAATCAGAAAAGGCTCGAAATGTCTAGAGGGG 120
CTGACTGATGCAGCGGTGACCCGGACTGGAGACAGTTGGA 160
CGCGATCATCTCTGGTGCTTTTGTTCAACCTTGAAACCTT 200
        210        220        230        240
....|....|....|....|....|....|....|....|....|
CGCCACAGGAGACTTGCCTGAGCAGAGAAGCAAACGTGGA 240
GAAACAAAGAGAGATCTAGCGAAAAGCCTCTGGGACCAAG 280
GAGGGGAGGTGGGACTCTGGGTTGGCGGTGGCACCTGCTG 320
CCGGCTATTAATAATAGGGTCGCGATGCGTTTATAAGGTG 360
TTTGATTAAACAAAGACTCTATGAGAGAAGAATAACTAGC 400
        410        420        430        440
....|....|....|....|....|....|....|....|....|
AACAGCCCCACGTCTGAGTCGTCGCCTCCGACCTTTTTCA 440
ACGTGGGTTCTTTGGGCCGAGCGTCGTTTGCCGAGAACTA 480
GATCTCACCTGACCCCAGACGCTGAAAACAAGCGCTGTGG 520
CATCCTGGGCCACCCAAGCTGACAAGGGCGCGCCCCCTGA 560
GCACACGAGGTGCCCCACGAGGGGGAGGGACCCACAGCCG 600
        610        620        630        640
....|....|....|....|....|....|....|....|....|
TCCCGCCCGCACCGCGGTGTCCGCTGCGGGCACCTGCAGC 640
CGAGCCGCCACCCGCAGTCGCAGCGCGTCCGGCGGCCGAA 680
CCCGGTCGTCAGCTCGTCAGCACCTGCTCTGCTTCTCTCC 720
CGCCCGCCGCCGCGCTGCACGCCTCGAGCGCTCCCTCGGC 760
CCCGGCGGGGACCGGGGACCCCGCAGCCACCGCCATGCTG 800
        810        820        830        840
....|....|....|....|....|....|....|....|....|
CCTGTGCTCTACACCGGCCTGGCGGGGCTGCTGCTGCTGC 840
CTCTGCTGCTCACCTGCTGCTGCCCCTACCTCCTCCAGGA 880
CGTGCGGTTCTTCCTGCAACTGGCCAACATGGCCCGGCAG 920
GTGCGCAGCTACCGGCAGCGGCGACCCGTGCGCACCATCC 960
TGCATGTCTTCTTGGAGCAAGCGCGCAAGACCCCGCACAA 1000
        1010       1020       1030       1040
....|....|....|....|....|....|....|....|....|
GCCCTTCCTGCTGTTTCGCGACGAGACGCTTACCTACGCC 1040
CAGGTAGACCGGCGCAGCAACCAAGTAGCGCGAGCGCTGC 1080
ATGATCACCTGGGCCTGCGGCAGGGGGATTGCGTGGCCCT 1120
CTTCATGGGCAATGAGCCGGCCTACGTGTGGCTCTGGCTG 1160
GGACTGCTCAAACTGGGCTGTCCCATGGCGTGCCTCAACT 1200
```

Fig. 58A

```
        1210        1220        1230        1240
...|....|....|....|....|....|....|....|....|....|
ACAACATCCGTGCCAAGTCTCTGCTACACTGCTTTCAGTG 1240
CTGCGGGGCGAAGGTGCTGCTGGCCTCCCCAGAGCTACAC 1280
GAAGCTGTCGAGGAGGTTCTTCCAACCCTGAAAAAGGAGG 1320
GCGTGTCCGTCTTCTACGTAAGCAGAACTTCTAACACTAA 1360
TGGCGTGGACACAGTACTGGACAAAGTAGACGGGGTGTCG 1400
        1410        1420        1430        1440
...|....|....|....|....|....|....|....|....|....|
GCGGACCCCATCCCGGAGTCGTGGAGGTCTGAAGTCACGT 1440
TCACCACACCCGCAGTCTACATATATACTTCGGGCACCAC 1480
AGGTCTTCCAAAGGCTGCAACCATTAATCACCATCGCCTC 1520
TGGTATGGGACCAGCCTTGCCCTGAGGTCCGGAATTAAGG 1560
CTCATGACGTCATCTACACCACCATGCCCCTGTACCACAG 1600
        1610        1620        1630        1640
...|....|....|....|....|....|....|....|....|....|
CGCGGCGCTCATGATTGGCCTCCACGGATGCATTGTGGTT 1640
GGGGCTACATTTGCTTTGCGGAGCAAATTTTCAGCCAGCC 1680
AGTTTTGGGACGACTGCAGGAAATACAACGCCACTGTCAT 1720
TCAGTACATCGGTGAACTGCTTCGGTACCTCTGCAACACG 1760
CCCCAGAAACCAAATGACCGGGACCACAAAGTGAAAATAG 1800
        1810        1820        1830        1840
...|....|....|....|....|....|....|....|....|....|
CACTAGGAAATGGCTTACGAGGAGATGTGTGGAGAGAGTT 1840
CATCAAGAGATTTGGGGACATTCACATTTATGAGTTCTAC 1880
GCTTCCACTGAAGGCAACATTGGATTTATGAACTATCCAA 1920
GAAAAATCGGAGCTGTTGGAAGAGAAAATTACCTACAAAA 1960
AAAAGTTGTAAGGCACGAGCTGATCAAGTATGACGTGGAG 2000
        2010        2020        2030        2040
...|....|....|....|....|....|....|....|....|....|
AAGGATGAGCCTGTCCGTGATGCAAATGGATATTGCATCA 2040
AAGTCCCCAAAGGAGAGGTTGGACTCTTGATTTGCAAAAT 2080
CACAGAGCTCACACCATTTTTTGGCTATGCTGGAGGAAAG 2120
ACCCAGACAGAGAAGAAAAAGCTCAGAGATGTTTTTAAGA 2160
AAGGAGACGTCTACTTCAACAGTGGCGATCTCCTGATGAT 2200
        2210        2220        2230        2240
...|....|....|....|....|....|....|....|....|....|
CGACCGTGAAAATTTCATCTATTTTCACGACAGAGTTGGA 2240
GACACCTTCCGGTGGAAAGGAGAGAATGTAGCTACCACGG 2280
AAGTCGCTGACATTGTGGGACTGGTAGATTTTGTTGAAGA 2320
AGTGAATGTTTACGGTGTGCCCGTGCCAGGTCATGAAGGT 2360
CGCATCGGGATGGCCTCGATCAAGATGAAAGAAAACTACG 2400
        2410        2420        2430        2440
...|....|....|....|....|....|....|....|....|....|
AGTTCAATGGAAAGAAACTCTTTCAGCACATCTCGGAGTA 2440
CCTGCCCAGTTACTCGAGGCCTCGGTTCCTGAGAATACAA 2480
GATACCATTGAGATCACCGGGACTTTTAAACACCGCAAAG 2520
TGACCCTGATGGAAGAGGGCTTTAACCCCTCAGTCATCAA 2560
AGATACCTTGTATTTCATGGATGACACAGAAAAAACATAC 2600
```

Fig. 58B

mVLACS(FATP2)full length.DNA

```
        2610      2620      2630      2640
GTGCCCATGACTGAGGACATTTATAATGCCATAATTGATA 2640
AGACTCTGAAGCTCTGAATGTTGCCTGGCTCCTAACACTT 2680
CCAGAAAGAAACACAATAGGCCTAGCATAGCCCCTTCACA 2720
TGTGTAATCCAACTTTAACTTGATTAAAGGTTATAGGTGT 2760
GATTTTTCCTAGGAAATTATTCATTTAAAGGACAATTGTT 2800
        2810      2820      2830      2840
TGTTTGTTTGTTTGTTTTTTATTAATTACACCAGAACGTT 2840
TGCAAGTAAAAAGATTTAAAGTCACTTATTTTTCAATGTG 2880
CACCTGCCATTTGTCCTTGCAAACTTAGCTTCTTGGAGAG 2920
AGGGCCTTATTTTTTTAAAGACATAATAAACTATGTAAAC 2960
ACT 2963
```

Fig. 58C

mVLACS(FATP2)full length.prot

```
        10         20         30         40
MLPVLYTGLAGLLLLPLLLTCCCPYLLQDVRFFLQLANMA 40
RQVRSYRQRRPVRTILHVFLEQARKTPHKPFLLFRDETLT 80
YAQVDRRSNQVARALHDHLGLRQGDCVALFMGNEPAYVWL 120
WLGLLKLGCPMACLNYNIRAKSLLHCFQCCGAKVLLASPE 160
LHEAVEEVLPTLKKEGVSVFYVSRTSNTNGVDTVLDKVDG 200
        210        220        230        240
VSADPIPESWRSEVTFTTPAVYIYTSGTTGLPKAATINHH 240
RLWYGTSLALRSGIKAHDVIYTTMPLYHSAALMIGLHGCI 280
VVGATFALRSKFSASQFWDDCRKYNATVIQYIGELLRYLC 320
NTPQKPNDRDHKVKIALGNGLRGDVWREFIKRFGDIHIYE 360
FYASTEGNIGFMNYPRKIGAVGRENYLQKKVVRHELIKYD 400
        410        420        430        440
VEKDEPVRDANGYCIKVPKGEVGLLICKITELTPFFGYAG 440
GKTQTEKKKLRDVFKKGDVYFNSGDLLMIDRENFIYFHDR 480
VGDTFRWKGENVATTEVADIVGLVDFVEEVNVYGVPVPGH 520
EGRIGMASIKMKENYEFNGKKLFQHISEYLPSYSRPRFLR 560
IQDTIEITGTFKHRKVTLMEEGFNPSVIKDTLYFMDDTEK 600
        610        620        630        640
TYVPMTEDIYNAIIDKTLKL. 621
```

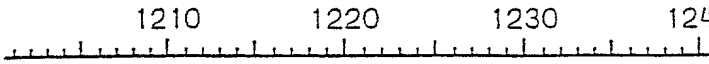
Fig. 59

mFATP4 partial.DNA

```
1        10        20        30        40
GATCAGCTCTTCTATATCTACACGTCGGGCACCACGGGGC 40
TACCCAAAGCTGCCATTGTGGTGCACAGCAGGTATTACCG 80
AATGGCTGCCCTGGTGTACTATGGATTCCGCATGCGGCCT 120
GATGACATTGTCTATGACTGCCTCCCCCTCTACCACTCAG 160
CAGGAAACATTGTGGGGATTGGCCAGTGCGTACTCCACGG 200
         210       220       230       240
CATGACTGTGGTGATCCGGAAGAAGTTTTCAGCCTCCCGG 240
TTCTGGGATGACTGTATCAAGTACAACTGCACAATTGTAC 280
AGTACATTGGTGAGCTTTGCCGCTACCTCCTGAACCAGCC 320
ACCCCGTGAGGCTGAGTCTCGGCACAAGGTGCGCATGGCA 360
CTGGGCAACGGTCTCCGGCAGTCCATCTGGACCGACTTCT 400
         410       420       430       440
CCAGCCGTTTCCACATTCCCAAGGTGGCCGAGTTCTACGG 440
GGCCACCGAGTGCAACTGTAGCTTGGGCAACTTTGACAGC 480
CAGGTGGGGGCCTGTGGCTTCAATAGCCGCATCCTGTCCT 520
TTGTGTACCCCATCCGCTTGGTACGAGTCAATGAGGATAC 560
CATGGAACTGATCCGGGGACCCGATGGCGTCTGCATTCCC 600
         610       620       630       640
TGTCAACCAGGCCAGCCAGGCCAGCTGGTGGGTCGCATCA 640
TCCAGCAGGACCCCCTACGCCGTTTTGATGGCTACCTCAA 680
CCAGGGTGCCAACAACAAGAAGATTGCTAGTGATGTCTTC 720
AAGAAAGGGGACCAAGCCTACCTCACTGGTGACGTGCTGG 760
TGATGGATGAGCTGGGCTACCTGTACTTCCGAGACCGCAC 800
         810       820       830       840
AGGGGACACGTTCCGCTGGAAAGGGGAGAATGTGTCTACC 840
ACTGAAGTGGAGGGCACACTCAGCCGCCTGCTTCAGATGG 880
CAGATGTGGCTGTTTATGGTGTTGAGGTGCCAGGAGCTGA 920
GGGCCGAGCAGGAATGGCTGCTGTGGCAAGCCCCACTAGC 960
AACTGTGACCTGGAGAGCTTTGCACAGACCTTGAAAAAGG 1000
         1010      1020      1030      1040
AGCTGCCCCTGTACGCCCGCCCCATCTTCCTCCGCTTCTT 1040
GCCTGAGCTGCACAAAACAGGAACCTTCAAGTTCCAGAAG 1080
ACAGAGTTGCGGAAGGAGGGCTTTGACCCGTCTGTTGTGA 1120
AAGACCCACTCTTCTATTTGGATGCCCGGACAGGCTGCTA 1160
TGTTGCACTGGACCAAGAGGCCTATACCCGCATCCAGGCA 1200
```

Fig. 60 A

mFATP4 partial.DNA



GGCGAGGAGAAGCTGTGATTTCCCCCACATCCCTCTGAGG 1240
GCCAGAGGATGCTGGATTCAGAGCCCCAGCTTCCACTCCA 1280
GAAGGGGTCTGGGCAAGGCCAGACCAAAGCTAGCAGGGCC 1320
CGCACCTTCACCCTAGGTGCTGATCCCCCT 1350

Fig. 60B

mFATP4partial.DNA

```
            10        20        30        40
DQLFYIYTSGTTGLPKAAIVVHSRYYRMAALVYYGFRMRP 40
DDIVYDCLPLYHSAGNIVGIGQCVLHGMTVVIRKKFSASR 80
FWDDCIKYNCTIVQYIGELCRYLLNQPPREAESRHKVRMA 120
LGNGLRQSIWTDFSSRFHIPKVAEFYGATECNCSLGNFDS 160
QVGACGFNSRILSFVYPIRLVRVNEDTMELIRGPDGVCIP 200
           210       220       230       240
CQPGQPGQLVGRIIQQDPLRRFDGYLNQGANNKKIASDVF 240
KKGDQAYLTGDVLVMDELGYLYFRDRTGDTFRWKGENVST 280
TEVEGTLSRLLQMADVAVYGVEVPGAEGRAGMAAVASPTS 320
NCDLESFAQTLKKELPLYARPIFLRFLPELHKTGTFKFQK 360
TELRKEGFDPSVVKDPLFYLDARTGCYVALDQEAYTRIQA 400
           410       420       430       440
GEEKL.  406
```

Fig. 61

mmFATP1 full length.DNA

```
         10         20         30         40
ATGCGGGCTCCTGGAGCAGGAACAGCCTCTGTGGCCTCAC 40
TGGCGCTGCTTTGGTTTCTGGGACTTCCGTGGACCTGGAG 80
CGCGGCGGCGGCGTTCTGTGTGTACGTGGGTGGCGGCGGC 120
TGGCGCTTTCTGCGTATCGTCTGCAAGACGGCGAGGCGAG 160
ACCTCTTTGGCCTCTCTGTTCTGATTCGTGTTCGGCTAGA 200
        210        220        230        240
GCTGCGACGACACCGGCGAGCAGGAGACACGATCCCGTGC 240
ATCTTCCAGGCTGTGGCCCGGCGACAACCAGAGCGCCTGG 280
CACTGGTGGACGCCAGTAGTGGTATATGCTGGACCTTCGC 320
ACAGCTGGACACCTACTCCAATGCTGTAGCCAACCTGTTC 360
CGCCAGCTGGGCTTTGCACCAGGCGATGTGGTGGCTGTGT 400
        410        420        430        440
TCCTGGAGGGCCGGCCGGAGTTCGTGGGACTGTGGCTGGG 440
CCTGGCCAAGGCCGGTGTGGTGGCTGCTCTTCTCAATGTC 480
AACCTGAGGCGGGAGCCCCTGGCCTTCTGCCTGGGCACAT 520
CAGCTGCCAAGGCCCTCATTTATGGCGGGGAGATGGCAGC 560
GGCGGTGGCGGAGGTGAGCGAGCAGCTGGGGAAGAGCCTC 600
        610        620        630        640
CTCAAGTTCTGCTCTGGAGATCTGGGGCCTGAGAGCATCC 640
TGCCTGACACGCAGCTCCTGGACCCCATGCTTGCTGAGGC 680
GCCCACCACACCCCTGGCACAAGCCCCAGGCAAGGGCATG 720
GATGATCGGCTGTTTTACATCTATACTTCTGGGACCACCG 760
GGCTTCCTAAGGCTGCCATTGTGGTGCACAGCAGGTACTA 800
        810        820        830        840
CCGCATTGCTGCCTTTGGCCACCATTCCTACAGCATGCGT 840
GCCGCCGATGTGCTCTATGACTGCCTGCCACTCTACCACT 880
CTGCAGGGAACATCATGGGTGTGGGGCAGTGCGTCATCTA 920
CGGGTTGACGGTGGTACTGCGCAAGAAGTTCTCCGCCAGC 960
CGCTTCTGGGATGACTGTGTCAAGTACAATTGCACGGTAG 1000
        1010       1020       1030       1040
TGGATGACATAGGTGAAATCTGCCGCTACCTGCTGAGGCA 1040
GCCGGTTCGCGACGTGGAGCAGCGACACCGCGTGCGCCTG 1080
GCCGTGGGTAATGGGCTGCGGCCAGCCATCTGGGAGGAGT 1120
TCACGCAGCGCTTCGGTGTGCCACAGATCGGCGAGTTCTA 1160
CGGCGCTACCGAGTGCAACTGCAGCATTGCCAACATGGAC 1200
```

Fig. 62A

```
        1210      1220      1230      1240
GGCAAGGTCGGCTCCTGCGGCTTCAACAGCCGTATCCTCA 1240
CGCATGTGTACCCCATCCGTCTGGTCAAGGTCAATGAGGA 1280
CACGATGGAGCCACTGCGGGACTCCGAGGGCCTCTGCATC 1320
CCGTGCCAGCCCGGGGAACCCGGCCTTCTCGTGGCCAGA 1360
TCAACCAGCAGGACCCTCTGCGGCGTTTCGATGGTTATGT 1400
        1410      1420      1430      1440
TAGTGACAGTGCCACCAACAAGAAGATTGCCCACAGCGTT 1440
TTCCGAAAGGGCGATAGCGCCTACCTCTCAGGTGACGTGC 1480
TAGTGATGGACGAGCTGGGCTACATGTATTTCCGTGACCG 1520
CAGCGGGGACACCTTCCGCTGGCGCGGGGAGAACGTGTCC 1560
ACCACGGAGGTGGAAGCCGTGCTGAGCCGCCTACTGGGCC 1600
        1610      1620      1630      1640
AGACGGACGTGGCTGTGTATGGGGTGGCTGTGCCAGGAGT 1640
GGAGGGGAAAGCTGGCATGGCAGCCATCGCAGATCCCCAC 1680
AGCCAGTTGGACCCTAACTCAATGTACCAGGAATTACAGA 1720
AGGTTCTTGCATCCTATGCTCGGCCCATCTTCCTGCGTCT 1760
TCTGCCCCAGGTGGATACCACAGGCACCTTCAAGATCCAG 1800
        1810      1820      1830      1840
AAGACCCGGCTGCAGCGTGAAGGCTTTGACCCCCGTCAGA 1840
CCTCAGACAGGCTCTTCTTTCTAGACCTGAAGTCCGGCAC 1880
GAGGTATCTACCCCTGGATGAGAGAGTCCATGCCCGCATT 1920
TGCGCAGGCGACTTCTCACTCTGAGCCTGGAGAGTGGGCT 1960
GGGCCTGGACTCCTGAGACCTGGGAGCCTGACACCCCTCT 2000
        2010      2020      2030      2040
TCGGGTGCTTCTCCTGCCTGGCCACATGGACAGCAGCACC 2040
TGTGAGAGTAGGAAAATGGAACCTGAGTGGCTGGGACCCC 2080
TCTCCTACTTCCCACTATGCATCCATTTTGCCTCTGCCTT 2120
GATCTTTTTCTCCATCTCTTTTCTCCCTACCCAGCAGGAG 2160
CCCCACAAACACATGTTGGCTGCTGTGTCCTGCAGTTGGA 2200
        2210      2220      2230      2240
CCAGTGTCCAGGGGTACAGGCTTCAGGCTGTGACCCACAC 2240
TGGTACCCACCTCCCTTTCCTATTTTGCCTTAGGTTCATC 2280
CACGGTTCCCCTGTGGAGCAAGTGGGGGCCCACATAGCTG 2320
CTGTCCCTGCTGAGGGTTGGTAGCAATCACACCCTCATGT 2360
CAGCTGGGAGACACGCGCAGTCTCCCACTGACCCCCAATC 2400
        2410      2420      2430      2440
AACTGAAAATATTGTTTTGACTACTTTTTGTTTTTTTGTT 2440
TTTTTGTTTTTTTTTTTTTTCGAGACAGAGTTTCTCTGTA 2480
TAGCCCTGGCTGTCCTGGAACTCACTTTGTAGACCAGGCT 2520
GGCCTCGAACTCAAAAATCCTCCTGACTCTGCCTCTGCTT 2560
CCCAAGTGCTGGGATTAAAGACGTGCGCCACCACCGCCTG 2600
```

Fig. 62B

mmFATP1 full length.DNA

```
        2610        2620        2630        2640
GCTGTTTTGTATTTTTGTTTTGTTTTGACGATAGGGTCTC 2640
ACTGTGGAGGCCAAGCTGGCCTCAGACTCCCCACCCCATT 2680
GCCTCTGGGCACCATTCTATATTCTCAGACTGATGACAAT 2720
GCACTAGTGTCCCTAGGAGTCTTGAGTCTGCACTTTCCCC 2760
TCATAGCCTCAAGCTTCCAGAACTGACTCTGATCACTTGG 2800
        2810        2820        2830        2840
ATGTGGCTAGTGTTGGCTCTACCCACATGTGTCAATTCAG 2840
GGGTCCCCAGGCATAGTCTCTGGAAGCCCTCACCCGGAAA 2880
AAGCTTGGAGAGACCCAGGAAGGTTGTTGTGTTCTCTTGG 2920
GCACCCCCTGGTGGCAGTCCTGGGCATGCTTCCGCACTGT 2960
ACTGGTGCATATAGCCCAGACCTATGACATTTGAGGTCTA 3000
        3010        3020        3030        3040
CCCTTCTGGCTCCTGTGGTCCCCATTGAGATCCTTGGTGA 3040
CTCACCTCAGTCACCAAGCAGAGCCTCTGCCTGCCTTCAT 3080
CTTCAAGGTCATGAAGGATGTGGACAGAGCAGCTACAGGC 3120
TGCCAGCAGTCAACCACATGAGAGTGTTACTTCCTTGTTG 3160
GTTTTTAAAAAATAAATGTGCTGAGCCTCGAAAAAAAAAA 3200
        3210        3220        3230        3240
AAAAAAAAAAAAAAAAA 3217
```

Fig. 62C

mmFATP1 full length.protein

```
          10         20         30         40
..........|..........|..........|..........|
MRAPGAGTASVASLALLWFLGLPWTWSAAAAFCVYVGGGG 40
WRFLRIVCKTARRDLFGLSVLIRVRLELRRHRRAGDTIPC 80
IFQAVARRQPERLALVDASSGICWTFAQLDTYSNAVANLF 120
RQLGFAPGDVVAVFLEGRPEFVGLWLGLAKAGVVAALLNV 160
NLRREPLAFCLGTSAAKALIYGGEMAAAVAEVSEQLGKSL 200
        210        220        230        240
..........|..........|..........|..........|
LKFCSGDLGPESILPDTQLLDPMLAEAPTTPLAQAPGKGM 240
DDRLFYIYTSGTTGLPKAAIVVHSRYYRIAAFGHHSYSMR 280
AADVLYDCLPLYHSAGNIMGVGQCVIYGLTVVLRKKFSAS 320
RFWDDCVKYNCTVVDDIGEICRYLLRQPVRDVEQRHRVRL 360
AVGNGLRPAIWEEFTQRFGVPQIGEFYGATECNCSIANMD 400
        410        420        430        440
..........|..........|..........|..........|
GKVGSCGFNSRILTHVYPIRLVKVNEDTMEPLRDSEGLCI 440
PCQPGEPGLLVGQINQQDPLRRFDGYVSDSATNKKIAHSV 480
FRKGDSAYLSGDVLVMDELGYMYFRDRSGDTFRWRGENVS 520
TTEVEAVLSRLLGQTDVAVYGVAVPGVEGKAGMAAIADPH 560
SQLDPNSMYQELQKVLASYARPIFLRLLPQVDTTGTFKIQ 600
        610        620        630        640
..........|..........|..........|..........|
KTRLQREGFDPRQTSDRLFFLDLKSGTRYLPLDERVHARI 640
CAGDFSL 647
```

Fig. 63

```
        10         20         30         40
GGGCGGAGGCCGAGCCCAGTCGCCAGCTCCTGCTCTGCTC 40
CTCTCCCGCCTGCCGCCGCGCTGCACGCCTCGAGCACTCC 80
CTCGGCCCCGGCGGGGACCGGGGACCCCGCAGCTACCGCC 120
ATGCTGCCAGTGCTCTACACCGGCCTGGCGGGGCTGCTGC 160
TGCTGCCTCTGCTGCTCACCTGCTGCTGCCCCTACCTCCT 200
        210        220        230        240
CCAAGATGTGCGGTACTTCCTGCGGCTGGCCAACATGGCC 240
CGGCGGGTGCGCAGCTACCGGCAGCGGCGACCCGTGCGTA 280
CCATCCTGCGGGCCTTCCTGGAACAAGCGCGCAAGACCCC 320
ACACAAGCCCTTCCTGCTGTTCCGAGACGAGACGCTCACC 360
TACGCCCAGGTGGACCGGCGCAGCAACCAAGTGGCGCGGG.400
        410        420        430        440
CGCTGCACGATCAACTGGGCCTACGACAGGGGGATTGCGT 440
AGCCCTCTTCATGGGCAATGAGCCGGCCTACGTGTGGATC 480
TGGCTGGGACTGCTCAAACTGGGCTGTCCCATGGCGTGCC 520
TCAACTACAACATTCGTGCCAAGTCTCTGCTGCACTGCTT 560
TCAATGCTGCGGGGCGAAGGTGCTGCTGGCCTCCCCAGAT 600
        610        620        630        640
CTACAAGAAGCTGTGGAGGAGGTTCTTCCAACCCTGAAAA 640
AGGATGCCGTGTCCGTCTTTTACGTAAGCAGAACTTCTAA 680
CACAAATGGTGTGGACACAATACTGGACAAAGTAGACGGA 720
GTGTCGGCGGAACCCACCCCGGAGTCGTGGAGGTCTGAAG 760
TCACTTTTACCACGCCAGCAGTATACATTTATACTTCGGG 800
        810        820        830        840
AACCACAGGTCTTCCAAAAAGCGGAACCATCAATCATCAT 840
CGCCTAAGGTATGGGACAAGCCTTGCTATGTCGAGTGGGA 880
ATCACGGCCAAGGATGTCATCTATACCAACAATGCCCCTG 920
TTCCAACAGTGCAACGCTCAAGATCGGCCTTCACGGATGC 960
ATCCTGGGTTGGGGCTACTTTAACCTTGGCGGGGCAAATT 1000
        1010       1020       1030       1040
CTCAAGCAAGCCAATTTTGGGAACGACTGGCAGGAAATAC 1040
AACGTCAACGGTCATTCAGTACATTGGTGAACTGCTTCGG 1080
TACCTGTGCAACACACCGCAGAAACCAAATGACCGGGACC 1120
ACAAAGTGAAAAAAGCCCTGGGAAATGGCTTACGAGGAGA 1160
TGTGTGGAGAGAGTTCATCAAGAGATTTGGGGACATCCAC 1200
```

Fig. 64A

```
        1210        1220        1230        1240
GTGTATGAGTTCTACGCATCCACTGAAGGCAACATTGGAT 1240
TTGTGAACTATCCAAGGAAAATCGGTGCTGTCGGGAGAGC 1280
AAACTACCTACAAAGAAAAGTTGCAAGGTATGAGCTGATC 1320
AAGTATGACGTGGAGAAGGACGAGCCGGTCCGTGACGCAA 1360
ATGGATATTGCATCAAAGTCCCCAAAGGTGAGGTTGGACT 1400
        1410        1420        1430        1440
CTTGGTTTGCAAAATCACACAGCTCACACCATTTATTGGC 1440
TATGCTGGAGGAAAGACCCAGACAGAGAAGAAAAAACTCA 1480
GAGATGTCTTTAAGAAAGGCGACATCTACTTCAACAGCGG 1520
AGACCTCCTGATGATCGACCGTGAGAACTTCGTCTACTTT 1560
CACGACAGGGTTGGAGATACTTTCCGGTGGAAAGGAGAGA 1600
        1610        1620        1630        1640
ACGTAGCTACCACAGAAGTCGCTGACATCGTGGGACTGGT 1640
AGATTTTGTTGAAGAAGTGAATGTGTATGGCGTGCCTGTG 1680
CCAGGTCATGAGGGTCGAATTGGGATGGCCTCCCTCAAGA 1720
TCAAAGAAAACTACGAGTTCAATGGAAAGAAACTCTTTCA 1760
ACACATCGCGGAGTACCTGCCCAGTTACGCGAGGCCTCGG 1800
        1810        1820        1830        1840
TTCCTGAGGATACAAGATACCATTGAGATCACTGGGACTT 1840
TTAAACACCGCAAAGTGACCCTGATGGAAGAGGGCTTCAA 1880
TCCCACAGTCATCAAAGATACCTTGTATTTCATGGATGAT 1920
GCAGAGAAAACATTTGTGCCCATGACTGAGAACATTTATA 1960
ATGCCATAATTGATAAAACTCTGAAGCTCTGAATATTCCC 2000
        2010        2020        2030        2040
TGGTGGTTTAGCTCATGACATTTCCAGAAAGAAACTCGAT 2040
AGACCTCGCAGAGCCACTTCATACGTAGAATCCAACTTTA 2080
ACTTGATTGAAGACTATAAGGTGCGATTTTATTTTTAGGA 2120
AATTATTCATTAAAAGGATAGTTTTTTTTTTTTTTTTAA 2160
TTACACCTGAACCTTTGCAAGTAAAAAGATTTAGAGACAA 2200
        2210        2220        2230        2240
TTATTTTTCAATGTGCACCTGCCATTTGTCCTTGCAAACT 2240
AAGCTTCTTGGAGAGAGGGCCTTATTTTTTTAAAGACATA 2280
ATAAACTATATTAACACTAAAAAAAAAAAAAAAAAAAAA 2320
AAAAAAAAAAAAAAAAAA 2338
```

Fig. 64B

mmFATP2 full length.protein

```
         10        20        30        40
MLPVLYTGLAGLLLLPLLLTCCCPYLLQDVRYFLRLANMA 40
RRVRSYRQRRPVRTILRAFLEQARKTPHKPFLLFRDETLT 80
YAQVDRRSNQVARALHDQLGLRQGDCVALFMGNEPAYVWI 120
WLGLLKLGCPMACLNYNIRAKSLLHCFQCCGAKVLLASPD 160
LQEAVEEVLPTLKKDAVSVFYVSRTSNTNGVDTILDKVDG 200
        210       220       230       240
VSAEPTPESWRSEVTFTTPAVYIYTSGTTGLPKSGTINHH 240
RLRYGTSLAMSSGNHGQGCHLYQQCPCSNSATLKIGLHGC 280
ILGWGYFNLGGANSQASQFWERLAGNTTSTVIQYIGELLR 320
YLCNTPQKPNDRDHKVKKALGNGLRGDVWREFIKRFGDIH 360
VYEFYASTEGNIGFVNYPRKIGAVGRANYLQRKVARYELI 400
        410       420       430       440
KYDVEKDEPVRDANGYCIKVPKGEVGLLVCKITQLTPFIG 440
YAGGKTQTEKKKLRDVFKKGDIYFNSGDLLMIDRENFVYF 480
HDRVGDTFRWKGENVATTEVADIVGLVDFVEEVNVYGVPV 520
PGHEGRIGMASLKIKENYEFNGKKLFQHIAEYLPSYARPR 560
FLRIQDTIEITGTFKHRKVTLMEEGFNPTVIKDTLYFMDD 600
        610       620       630       640
AEKTFVPMTENIYNAIIDKTLKL. 624
```

Fig. 65

```
      10         20         30         40
GAAAGCTCTGAGAGCGGGTGCAGTCTGGCCTGGCGTCTCG 40
CGTACCTGGCCCGGGAGCAGCCGACACACACCTTCCTCAT 80
CCACGGCGCGCAGCGCTTTAGCTACGCGGAGGCTGAGCGC 120
GAGAGCAACCGGATTGCTCGCGCCTTTCTGCGCGCACGGG 160
GCTGGACCGGGGGCCGCCGAGGCTCGGGCAGGGGCAGCAC 200
      210        220        230        240
TGAGGAAGGCGCACGCGTGGCGCCTCCGGCTGGAGATGCG 240
GCTGCTAGAGGGACGACCGCGCCCCCTCTGGCACCCGGGG 280
CGACCGTGGCGCTGCTCCTCCCAGCGGGCCCGGATTTCCT 320
TTGGATTTGGTTCGGACTGGCCAAAGCTGGCCTGCGCACG 360
GCCTTTGTGCCCACCGCTTTACGCCGAGGACCCCTGCTGC 400
      410        420        430        440
ACTGCCTCCGCAGCTGCGGTGCGAGTGCGCTCGTGCTGGC 440
CACAGAGTTCCTGGAGTCCCTGGAGCCGGACCTGCCGGCC 480
TTGAGAGCCATGGGGCTCCACCTATGGGCGACGGGCCCTG 520
AAACTAATGTAGCTGGAATCAGCAATTTGCTATCGGAAGC 560
AGCAGACCAAGTGGATGAGCCAGTGCCGGGGTACCTCTCT 600
      610        620        630        640
GCCCCCCAGAACATAATGGACACCTGCCTGTACATCTTCA 640
CCTCTGGCACTACTGGCCTGCCCAAGGCTGCTCGAATCAG 680
TCATCTGAAGGTTCTACAGTGCCAGGGATTCTACCATCTG 720
TGTGGAGTCCACCAGGAGGACGTGATCTACCTCGCACTCC 760
CACTGTACCACATGTCTGGCTCCCTTCTGGGCATTGTGGG 800
      810        820        830        840
CTGCTTGGGCATTGGGGCCACCGTGGTGCTGAAACCCAAG 840
TTCTCAGCTAGCCAGTTCTGGGACGATTGCCAGAAACACA 880
GGGTGACAGTGTTCCAGTACATTGGGGAGTTGTGCCGATA 920
CCTCGTCAACCAGCCCCCGAGCAAGGCAGAGTTTGACCAT 960
AAGGTGCGCTTGGCAGTGGGCAGTGGGTTGCGCCCAGACA 1000
      1010       1020       1030       1040
CCTGGGAGCGTTTCCTGCGGCGATTTGGACCTCTGCAGAT 1040
ACTGGAGACGTATGGCATGACAGAGGGCAACGTAGCTACG 1080
TTCAATTACACAGGACGGCAGGGTGCAGTGGGGCGAGCTT 1120
CCTGGCTTTACAAGCACATCTTCCCCTTCTCCTTGATTCG 1160
ATACGATGTCATGACAGGGGAGCCTATTCGGAATGCCCAG 1200
```

Fig. 66A

mmFATP3 partial.DNA

```
        1210        1220        1230        1240
GGGCACTGCATGACCACATCTCCAGGTGAGCCAGGCCTAC 1240
TGGTGGCCCCAGTGAGCCAGCAGTCCCCCTTCCTGGGCTA 1280
TGCTGGGGCTCCGGAGCTGGCCAAGGACAAGCTGCTGAAG 1320
GATGTCTTCTGGTCTGGGGACGTTTTCTTCAATACTGGGG 1360
ACCTCTTGGTCTGTGATGAGCAAGGCTTTCTTCACTTCCA 1400
        1410        1420        1430        1440
CGATCGTACTGGAGACACCATCAGGTGGAAGGGAGAGAAT 1440
GTGGCCACAACTGAAGTGGCTGAGGTCTTGGAGACCCTGG 1480
ACTTCCTTCAGGAGGTGAACATCTATGGAGTCACGGTGCC 1520
AGGGCACGAAGGCAGGGCAGGCATGGCGGCCTTGGCTCTG 1560
CGGCCCCCGCAGGCTCTGAACCTGGTGCAGCTCTACAGCC 1600
        1610        1620        1630        1640
ATGTTTCTGAGAACTTGCCACCGTATGCCCGACCTCGGTT 1640
TCTCAGGCTCCAGGAATCTTTGGCCACTACTGAGACCTTC 1680
AAACAGCAGAAGGTTAGGATGGCCAATGAGGGCTTTGACC 1720
CCAGTGTACTGTCTGACCCACTCTATGTTCTGGACCAAGA 1760
TATAGGGGCCTACCTGCCCCTCACACCTGCCCGGTACAGT 1800
        1810        1820        1830        1840
GCCCTCCTGTCTGGAGACCTTCGAATCTGAAACCTTCCAC 1840
TTGAGGGAGGGGCTCGGAGGGTACAGGCCACCATGGCTGC 1880
ACCAGGGAGGGTTTTCGGGTATCTTTTGTATATGGAGTCA 1920
TTATTTTGTAATAAACAGCTGGAGCTTAAAAAAAAAAAAA 1960
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1998
```

Fig. 66B

mmFATP3 partial.protein

```
         10        20        30        40
         |         |         |         |
ESSESGCSLAWRLAYLAREQPTHTFLIHGAQRFSYAEAER 40
ESNRIARAFLRARGWTGGRRGSGRGSTEEGARVAPPAGDA 80
AARGTTAPPLAPGATVALLLPAGPDFLWIWFGLAKAGLRT 120
AFVPTALRRGPLLHCLRSCGASALVLATEFLESLEPDLPA 160
LRAMGLHLWATGPETNVAGISNLLSEAADQVDEPVPGYLS 200
        210       220       230       240
         |         |         |         |
APQNIMDTCLYIFTSGTTGLPKAARISHLKVLQCQGFYHL 240
CGVHQEDVIYLALPLYHMSGSLLGIVGCLGIGATVVLKPK 280
FSASQFWDDCQKHRVTVFQYIGELCRYLVNQPPSKAEFDH 320
KVRLAVGSGLRPDTWERFLRRFGPLQILETYGMTEGNVAT 360
FNYTGRQGAVGRASWLYKHIFPFSLIRYDVMTGEPIRNAQ: 400
        410       420       430       440
         |         |         |         |
GHCMTTSPGEPGLLVAPVSQQSPFLGYAGAPELAKDKLLK 440
DVFWSGDVFFNTGDLLVCDEQGFLHFHDRTGDTIRWKGEN 480
VATTEVAEVLETLDFLQEVNIYGVTVPGHEGRAGMAALAL 520
RPPQALNLVQLYSHVSENLPPYARPRFLRLQESLATTETF 560
KQQKVRMANEGFDPSVLSDPLYVLDQDIGAYLPLTPARYS 600
        610       620       630       640
         |         |         |         |
ALLSGDLRI. 610
```

Fig. 67

```
         10        20        30        40
ATGCTGCTTGGAGCCTCTCTGGTGGGGGCGCTACTGTTCT 40
CCAAGCTAGTGCTGAAGCTGCCCTGGACCCAGGTGGGATT 80
CTCCCTGTTGCTCCTGTACTTGGGGTCTGGTGGCTGGCGT 120
TTCATCCGGGTCTTCATCAAGACGGTCAGGAGAGATATCT 160
TTGGTGGCATGGTGCTCCTGAAGGTGAAGACCAAGGTGCG 200
        210       220       230       240
ACGGTACCTTCAGGAGCGGAAGACGGTGCCCCTGCTGTTT 240
GCTTCAATGGTACAGCGCCACCCGGACAAGACAGCCCTGA 280
TTTTCGAGGGCACAGACACTCACTGGACCTTCCGCCAGCT 320
GGATGAGTACTCCAGTAGTGTGGCCAACTTCCTGCAGGCC 360
CGGGGCCTGGCCTCAGGCAATGTAGTTGCCCTCTTTATGG 400
        410       420       430       440
AAAACCGCAATGAGTTTGTGGGTCTGTGGCTAGGCATGGC 440
CAAGCTGGGCGTGGAGGCGGCTCTCATCAACACCAACCTT 480
AGGCGGGATGCCCTGCGCCACTGTCTTGACACCTCAAAGG 520
CACGAGCTCTCATCTTTGGCAGTGAGATGGCCTCAGCTAT 560
CTGTGAGATCCATGCTAGCCTGGAGCCCACACTCAGCCTC 600
        610       620       630       640
TTCTGCTCTGGATCCTGGGAGCCCAGCACAGTGCCCGTCA 640
GCACAGAGCATCTGGACCCTCTTCTGGAAGATGCCCCGAA 680
GCACCTGCCCAGTCACCCAGACAAGGGTTTTACAGATAAG 720
CTCTTCTACATCTACACATCGGGCACCACGGGGCTACCCA 760
AAGCTGCCATTGTGGTGCACAGCAGGTATTATCGTATGGC 800
        810       820       830       840
TTCCCTGGTGTACTATGGATTCCGCATGCGGCCTGATGAC 840
ATTGTCTATGACTGCCTCCCCCTCTACCACTCAAGCAGGA 880
AACATCGTGGGGATTGGCAGTGCTTACTCCACGGCATGAC 920
TGTGGTGATCCGGAAGAAGTTCTCAGCCTCCCGGTTCTGG 960
GATGATTGTATCAAGTACAACTGCACAGTGGTACAGTACA 1000
        1010      1020      1030      1040
TTGGCGAGCTCTGCCGCTACCTCCTGAACCAGCCACCCCG 1040
TGAGGCTGAGTCTCGGCACAAGGTGCGCATGGCACTGGGC 1080
AACGGTCTCCGGCAGTCCATCTGGACCGACTTCTCCAGCC 1120
GTTTCCACATCCCCCAGGTGGCTGAGTTCTATGGGGCCAC 1160
TGAATGCAACTGTAGCCTGGGCAACTTTGACAGCCGGGTG 1200
```

Fig. 68A

```
        1210        1220        1230        1240
GGGGCCTGTGGCTTCAATAGCCGCATCCTGTCCTTTGTGT 1240
ACCCTATCCGTTTGGTACGTGTCAATGAGGATACCATGGA 1280
ACTGATCCGGGGACCCGATGGAGTCTGCATTCCCTGTCAA 1320
CCAGGTCAGCCAGGCCAGCTGGTGGGTCGCATCATCCAGC 1360
AGGACCCTCTGCGCCGTTTCGACGGGTACCTCAACCAGGG 1400
        1410        1420        1430        1440
TGCCAACAACAAGAAGATTGCTAATGATGTCTTCAAGAAG 1440
GGGGACCAAGCCTACCTCACTGGTGACGTCCTGGTGATGG 1480
ATGAGCTGGGTTACCTGTACTTCCGAGATCGCACTGGGGA 1520
CACGTTCCGCTGGAAAGGGGAGAATGTATCTACCACTGAG 1560
GTGGAGGGCACACTCAGCCGCCTGCTTCATATGGCAGATG 1600
        1610        1620        1630        1640
TGGCAGTTTATGGTGTTGAGGTGCCAGGAACTGAAGGCCG 1640
AGCAGGAATGGCTGCCGTTGCAAGTCCCATCAGCAACTGT 1680
GACCTGGAGAGCTTTGCACAGACCTTGAAAAAGGAGCTGC 1720
CTCTGTATGCCCGCCCCATCTTCCTGCGCTTCTTGCCTGA 1760
GCTGCACAAGACAGGGACCTTCAAGTTCCAGAAGACAGAG 1800
        1810        1820        1830        1840
TTGCGGAAGGAGGGCTTTGACCCATCTGTTGTGAAAGACC 1840
CGCTGTTCTATCTGGATGCTCGGAAGGGCTGCTACGTTGC 1880
ACTGGACCAGGAGGCCTATACCCGCATCCAGGCAGGCGAG 1920
GAGAAGCTGTGATTTCCCCCTACATCCCTCTGAGGGCCAG 1960
AAGATGCTGGATTCAGAGCCCTAGCGTCCACCCCAGAGGG 2000
        2010        2020        2030        2040
TCCTGGGCAATGCCAGACCAAAGCTAGCAGGGCCCGCACC 2040
TCCGCCCCTAGGTGCTGATCTCCCCTCTCCCAAACTGCCA 2080
AGTGACTCACTGCCGCTTCCCCGACCCTCCAGAGGCTTTC 2120
TGTGAAAGTCTCATCCAAGCTGTGTCTTCTGGTCCAGGCG 2160
TGGCCCCTGGCCCCAGGGTTTCTGATAGGCTCCTTTAGGA 2200
        2210        2220        2230        2240
TGGTATCTTGGGTCCAGCGGGCCAGGGTGTGGGAGAGGAG 2240
TCACTAAGATCCCTCCAATCAGAAGGGAGCTTACAAAGGA 2280
ACCAAGGCAAAGCCTGTAGACTCAGGAAGCTAAGTGGCCA 2320
GAGACTATAGTGGCCAGTCATCCCATGTCCACAGAGGATC 2360
TTGGTCCAGAGCTGCCAAAGTGTCACCTCTCCCTGCCTGC 2400
        2410        2420        2430        2440
ACCTCTGGGGAAAAGAGGACAGCATGTGGCCACTGGGCAC 2440
CTGTCTCAAGAAGTCAGGATCACACACTCAGTCCTTGTTT 2480
CTCCAGGTTCCCTTGTTCTTGTCTCGGGGAGGGAGGGACG 2520
AGTGTCCTGTCTGTCCTTCCTGCCTGTCTGTGAGTCTGTG 2560
TTGCTTCTCCATCTGTCCTAGCCTGAGTGTGGGTGGAACA 2600
```

Fig. 68B

mmFATP4 full length.DNA

```
        2610      2620      2630      2640
GGCATGAGGAGAGTGTGGCTCAGGGGCCAATAAACTCTGC 2640
CTTGACTCCTCTTAAAAAAAAAAAAAAAAAAAAAAAAAAA 2680
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2710
```

Fig. 68C

mmFATP4 full length.protein

```
           10        20        30        40
MLLGASLVGALLFSKLVLKLPWTQVGFSLLLLYLGSGGWR 40
FIRVFIKTVRRDIFGGMVLLKVKTKVRRYLQERKTVPLLF 80
ASMVQRHPDKTALIFEGTDTHWTFRQLDEYSSSVANFLQA 120
RGLASGNVVALFMENRNEFVGLWLGMAKLGVEAALINTNL 160
RRDALRHCLDTSKARALIFGSEMASAICEIHASLEPTLSL 200
          210       220       230       240
FCSGSWEPSTVPVSTEHLDPLLEDAPKHLPSHPDKGFTDK 240
LFYIYTSGTTGLPKAAIVVHSRYYRMASLVYYGFRMRPDD 280
IVYDCLPLYHSSRKHRGDWQCLLHGMTVVIRKKFSASRFW 320
DDCIKYNCTVVQYIGELCRYLLNQPPREAESRHKVRMALG 360
NGLRQSIWTDFSSRFHIPQVAEFYGATECNCSLGNFDSRV 400
          410       420       430       440
GACGFNSRILSFVYPIRLVRVNEDTMELIRGPDGVCIPCQ 440
PGQPGQLVGRIIQQDPLRRFDGYLNQGANNKKIANDVFKK 480
GDQAYLTGDVLVMDELGYLYFRDRTGDTFRWKGENVSTTE 520
VEGTLSRLLHMADVAVYGVEVPGTEGRAGMAAVASPISNC 560
DLESFAQTLKKELPLYARPIFLRFLPELHKTGTFKFQKTE 600
          610       620       630       640
LRKEGFDPSVVKDPLFYLDARKGCYVALDQEAYTRIQAGE 640
EKL. 644
```

Fig. 69

mmFATP5 full length.DNA

```
        10        20        30        40
CACTCATCAGAGCTAAGAGAGACTACACGCTCTCATCTAC 40
TTCAGAAAGAGCCAATGCCATGGGTATTTGGAAGAAACTA 80
ACCTTACTGCTGTTGCTGCTTCTGCTGGTTGGCCTGGGGC 120
AGCCCCCATGGCCAGCAGCTATGGCTCTGGCCCTGCGTTG 160
GTTCCTGGGAGACCCCACATGCCTTGTGCTGCTTGGCTTG 200
       210       220       230       240
GCATTGCTGGGCAGACCCTGGATCAGCTCCTGGATGCCCC 240
ACTGGCTGAGCCTGGTAGGAGCAGCTCTTACCTTATTCCT 280
ATTGCCTCTACAGCCACCCCCAGGGCTACGCTGGCTGCAT 320
AAAGATGTGGCTTTCACCTTCAAGATGCTTTTCTATGGCC 360
TAAAGTTCAGGCGACGCCTTAACAAACATCCTCCAGAGAC 400
       410       420       430       440
CTTTGTGGATGCTTTAGAGCGGCAAGCACTGGCATGGCCT 440
GACCGGGTGGCCTTGGTGTGTACTGGGTCTGAGGGCTCCT 480
CAATCACAAATAGCCAGCTGGATGCCAGGTCCTGTCAGGC 520
AGCATGGGTCCTGAAAGCAAAGCTGAAGGATGCCGTAATC 560
CAGAACACAAGAGATGCTGCTGCTATCTTAGTTCTCCCGT 600
       610       620       630       640
CCAAGACCATTTCTGCTTTGAGTGTGTTTCTGGGGTTGGC 640
CAAGTTGGGCTGCCCTGTGGCCTGGATCAATCCACACAGC 680
CGAGGGATGCCCTTGCTACACTCTGTACGGAGCTCTGGGG 720
CCAGTGTGCTGATTGTGGATCCAGACCTCCAGGAGAACCT 760
GGAAGAAGTCCTTCCCAAGCTGCTAGCTGAGAACATTCAC 800
       810       820       830       840
TGCTTCTACCTTGGCCACAGCTCACCCACCCCGGGAGTAG 840
AGGCTCTGGGAGCTTCCCTGGATGCTGCACCTTCTGACCC 880
AGTACCTGCCAGCCTTCGAGCTACGATTAAGTGGAAATCT 920
CCTGCCATATTCATCTTTACTTCAGGGACCACTGGACTCC 960
CAAAGCCAGCCATCTTATCACATGAGCGGGTCATACAAGT 1000
      1010      1020      1030      1040
GAGCAACGTGCTGTCCTTCTGTGGATGCAGAGCTGATGAT 1040
GTGGTCTATGACGTCCTACCTCTGTACCATACGATAGGGC 1080
TTGTCCTTGGATTCCTTGGCTGCTTACAAGTTGGAGCCAC 1120
CTGTGTCCTGGCCCCCAAGTTCTCTGCCTCCCGATTCTGG 1160
GCTGAGTGCCGGCAGCATGGCGTAACAGTGATCTTGTATG 1200
```

Fig. 70A

mmFATP5 full length.DNA

```
        1210      1220      1230      1240
TGGGTGAAATCCTGCGGTACTTGTGTAACGTCCCTGAGCA 1240
ACCAGAAGACAAGATACATACAGTGCGCTTGGCCATGGGA 1280
ACTGGACTTCGGGCAAATGTGTGGAAAAACTTCCAGCAAC 1320
GCTTTGGTCCCATTCGGATCTGGGAATTCTACGGATCCAC 1360
AGAGGGCAATGTGGGCTTAATGAACTATGTGGGCCACTGC 1400
        1410      1420      1430      1440
GGGGCTGTGGGAAGGACCAGCTGCATCCTTCGAATGCTGA 1440
CTCCCTTTGAGCTTGTACAGTTCGACATAGAGACAGCAGA 1480
GCCTCTGAGGGACAAACAGGGTTTTTGCATTCCTGTGGAG 1520
CCAGGAAAGCCAGGACTTCTTTTGACCAAGGTTCGAAAGA 1560
ACCAACCCTTCCTGGGCTACCGTGGTTCCCAGGCCGAGTC 1600
        1610      1620      1630      1640
CAATCGGAAACTTGTTGCGAATGTACGACGCGTAGGAGAC 1640
CTGTACTTCAACACTGGGGACGTGCTGACCTTGGACCAGG 1680
AAGGCTTCTTCTACTTTCAAGACCGCCTTGGTGACACCTT 1720
CCGGTGGAAGGGCGAAAACGTATCTACTGGAGAGGTGGAG 1760
TGTGTTTTGTCTAGCCTAGACTTCCTAGAGGAAGTCAATG 1800
        1810      1820      1830      1840
TCTATGGTGTGCCTGTGCCAGGGTGTGAGGGTAAGGTTGG 1840
CATGGCTGCTGTGAAACTGGCTCCTGGGAAGACTTTTGAT 1880
GGGCAGAAGCTATACCAGCATGTCCGCTCCTGGCTCCCTG 1920
CCTATGCCACACCTCATTTCATCCGTATCCAGGATTCCCT 1960
GGAGATCACAAACACCTACAAGCTGGTAAAGTCACGGCTG 2000
        2010      2020      2030      2040
GTGCGTGAGGGTTTTGATGTGGGGATCATTGCTGACCCCC 2040
TCTACATACTGGACAACAAGGCCCAGACCTTCCGGAGTCT 2080
GATGCCAGATGTGTACCAGGCTGTGTGTGAAGGAACCTGG 2120
AATCTCTGACCACCTAGCCAACTGGAAGGCAATCCAAAAG 2160
TGTAGAGATTGACACTAGTCAGCTTCACAAAGTTGTCCGG 2200
        2210      2220      2230      2240
GTTCCAGATGCCCATGGCCCAGTAGTACTTAGAGAATAAA 2240
CTTGAATGTGTATACAAAAAAAAAAAAAAAAAAAAA 2277
```

Fig. 70B

mmFATP5 full length.protein

```
         10         20         30         40
MGIWKKLTLLLLLLLLVGLGQPPWPAAMALALRWFLGDPT 40
CLVLLGLALLGRPWISSWMPHWLSLVGAALTLFLLPLQPP 80
PGLRWLHKDVAFTFKMLFYGLKFRRRLNKHPPETFVDALE 120
RQALAWPDRVALVCTGSEGSSITNSQLDARSCQAAWVLKA 160
KLKDAVIQNTRDAAAILVLPSKTISALSVFLGLAKLGCPV 200
        210        220        230        240
AWINPHSRGMPLLHSVRSSGASVLIVDPDLQENLEEVLPK 240
LLAENIHCFYLGHSSPTPGVEALGASLDAAPSDPVPASLR 280
ATIKWKSPAIFIFTSGTTGLPKPAILSHERVIQVSNVLSF 320
CGCRADDVVYDVLPLYHTIGLVLGFLGCLQVGATCVLAPK 360
FSASRFWAECRQHGVTVILYVGEILRYLCNVPEQPEDKIH 400
        410        420        430        440
TVRLAMGTGLRANVWKNFQQRFGPIRIWEFYGSTEGNVGL 440
MNYVGHCGAVGRTSCILRMLTPFELVQFDIETAEPLRDKQ 480
GFCIPVEPGKPGLLLTKVRKNQPFLGYRGSQAESNRKLVA 520
NVRRVGDLYFNTGDVLTLDQEGFFYFQDRLGDTFRWKGEN 560
VSTGEVECVLSSLDFLEEVNVYGVPVPGCEGKVGMAAVKL 600
        610        620        630        640
APGKTFDGQKLYQHVRSWLPAYATPHFIRIQDSLEITNTY 640
KLVKSRLVREGFDVGIIADPLYILDNKAQTFRSLMPDVYQ 680
AVCEGTWNL. 690
```

Fig. 71

```
         10        20        30        40
GCTCTCTGGGCCTATATCAAGCTGCTGAGGTACACGAAGC 40
GCCATGAGCGGCTCAACTACACGGTGGCGGACGTCTTCGA 80
ACGAAATGTTCAGGCCCATCCGGACAAGGTGGCTGTGGTC 120
AGTGAGACGCAACGCTGGACCTTCCGTCAGGTGAACGAGC 160
ATGCGAACAAGGTGGCCAATGTGCTGCAGGCTCAGGGCTA 200
        210       220       230       240
CAAAAAGGGCGATGTGGTGGCCCTGTTGCTGGAGAACCGC 240
GCCGAGTACGTGGCCACCTGGCTGGGTCTCTCCAAGATCG 280
GTGTGATCACACCGCTGATCAACACGAATCTGCGCGGTCC 320
CTCCCTGCTGCACAGCATCACGGTGGCCCATTGCTCGGCT 360
CTCATTTACGGCGAGGACTTCCTGGAAGCTGTCACCGACG 400
        410       420       430       440
TGGCCAAGGATCTGCCAGCGAACCTCACACTCTTCCAGTT 440
CAACAACGAGAACAACAACAGCGAGACGGAAAAGAACATA 480
CCGCAGGCCAAGAATCTGAACGCGCTGCTGACCACGGCCA 520
GCTATGAGAAGCCTAACAAGACGCAGGTTAACCACCACGA 560
CAAGCTGGTCTACATCTACACCTCCGGCACCACAGGATTG 600
        610       620       630       640
CCAAAGGCTGCGGTTATCTCTCACTCCCGTTATCTGTTTA 640
TCGCTGCTGGCATCCACTACACCATGGGTTTCCAGGAGGA 680
GGACATCTTCTACACGCCCTTGCCTTTGTACCACACCGCT 720
GGTGGCATTATGTGCATGGGTCAGTCGGTGCTCTTTGGCT 760
CCACGGTCTCCATTCGCAAGAAGTTCTCGGCATCCAACTA 800
        810       820       830       840
TTTCGCCGACTGCGCCAAGTATAATGCAACTATTGGTCAG 840
TATATCGGTGAGATGGCTCGCTACATTCTAGCTACGAAAC 880
CCTCGGAATACGACCAGAAACACCGAGTGCGTCTGGTCTT 920
TGGAAACGGACTGCGACCGCAGATTTGGCCACAGTTTGTG 960
CAGCGCTTCAACATTGCCAAGGTTGGCGAGTTCTACGGCG 1000
       1010      1020      1030      1040
CCACCGAGGGTAATGCGAACATCATGAATCATGACAACAC 1040
GGTGGGCGCCATCGGCTTTGTGTCGCGCATCCTGCCCAAG 1080
ATCTACCCAATCTCGATCATTCGCGCCGATCCGGACACCG 1120
GAGAGCCCATTAGAGATAGGAATGGCCTATGCCAACTGTG 1160
CGCTCCCAACGAGCCAGGCGTATTCATCGGCAAGATCGTC 1200
```

Fig. 72A

dmFATP partial.DNA

```
     1210      1220      1230      1240
AAAGGAAATCCTTCTCGCGAATTCCTCGGATACGTCGATG 1240
AAAAGGCCTCCGCGAAGAAGATTGTTAAGGATGTGTTCAA 1280
GCATGGCGATATGGCTTTCATCTCCGGAGATCTGCTGGTT 1320
GCCGACGAGAAGGGTTATCTGTACTTCAAGGATCGCACCG 1360
GTGACACCTTCCGCTGGAAGGGCGAGAATGTTTCCACCAG 1400
     1410      1420      1430      1440
CGAGGTGGAGGCGCAAGTCAGCAATGTGGCCGGTTACAAG 1440
GATACCGTCGTTTACGGCGTAACCATTCCGCACACCGAGG 1480
GAAGGGCCGGCATGGCCGCCATCTATGATCCGGAGCGAGA 1520
ATTGGACCTCGACGTCTTCGCCGCTAGCTTGGCCAAGGTG 1560
CTGCCCGCGTACGCTCGTCCCCAGATCATTCGATTGCTCA 1600
     1610      1620      1630      1640
CCAAGGTGGACCTGACTGGAACCTTTAAGCTGCGCAAGGT 1640
AGACCTGCAGAAGGAGGGCTACGATCCGAACGCGATCAAG 1680
GACGCGCTGTACTACCAGACTTCCAAGGGTCGGTACGAGC 1720
TGCTCACGCCCCAGGTTTACGACCAGGTGCAGCGCAACGA 1760
AATCCGCTTCTAAGAGCTGCAATAGAGTTGTGTCTGAACC 1800
     1810      1820      1830      1840
TTGCCTTTTGCCCAATATGCTGTTAATTAGTTTGTAAGGC 1840
TAAGTGTAGTAGAGGAAAATCGGGGGAAATCGGCAGCAAA 1880
GATCATTCAGCCTAGGAGAGATGCATCCGAAGCACATTTC 1920
CATGTCAACAATGCACTTTTGTATATCGTAAGCATATATA 1960
TATCGTATATCGTAAACGTAGTTGTATCTGCATTTGTGTA 2000
     2010      2020      2030      2040
GATGATAGCCTCCTATACGCATTTCAATTGTTTTTAGCGT 2040
GCTAAAGAACCTTGTTAAATGCAATTTCAGCTATTGTTTA 2080
GTCAGTTTTAGTGGCATTTACACTTCCATTCTCGTTGCGT 2120
TTCGTTTTTGCCTGTACATATGAGAAGCTCTGATGTTTTT 2160
GTATCAAATAAAGTTTTTTCCTTCACCACGGACCACGTGA 2200
     2210      2220      2230      2240
AAAAAAAAAAAAAAAAAAAAA 2221
```

Fig. 72B

dmFATP partial.protein

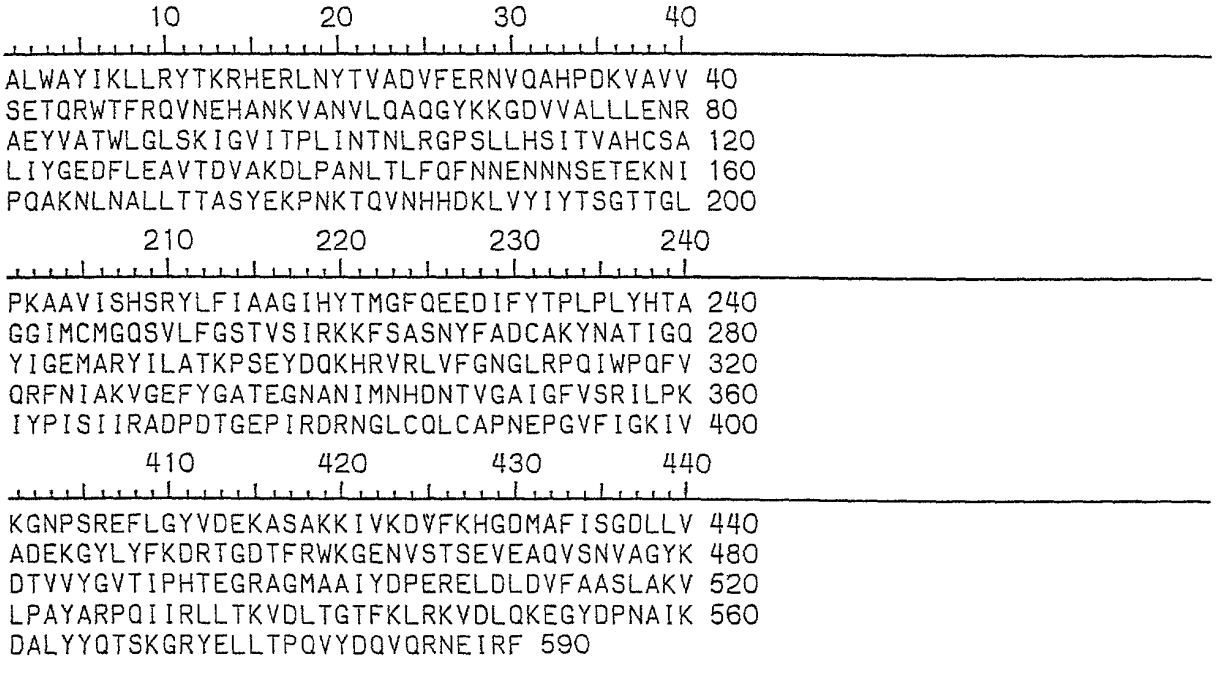
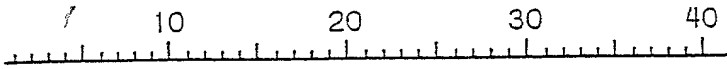


Fig. 73



AGTGTAGATACCACAGGAACGTTTAAAATCCAGAAGACCA 40
GACTGCAAAGGGAAGGATACGATCCACGGCTCACAACTGA 80
CCAGATCTACTTCCTAAACTCCAGAGCAGGGCGTTACGAG 120
CTTGTCAACGAGGAGCTGTACAATGCATTTGAACAAGGGC 160
AGGATTTCCCTTT 173

Fig. 74

drFATP partial.protein

```
         10        20        30        40
SVDTTGTFKIQKTRLQREGYDPRLTTDQIYFLNSRAGRYE 40
LVNEELYNAFEQGQDFP 57
```

Fig. 75

10 20 30 40

ATGAAGCTGGAGGAGCTTGTGACAGTTATGCTTCTCACAG 40
TGGCTGTCATTGCTCAGAATCTTCCGATTGGAGTAATATT 80
GGCTGGAGTTCTTATTTTATACATCACAGTGGTTCATGGA 120
GATTTCATTTATAGAAGTTATCTTACGTTGAATAGGGATT 160
TAACAGGATTGGCTCTAATTATTGAAGTCAAAATCGACCT 200

210 220 230 240

ATGGTGGAGGTTGCATCAGAATAAAGGAATCCATGAACTG 240
TTTTTGGATATTGTGAAAAAGAATCCAAATAAGCCGGCGA 280
TGATTGACATCGAGACGAATACAACAGAAACATACGCAGA 320
GTTCAATGCACATTGTAATAGATATGCCAATTATTTCCAG 360
GGTCTTGGCTATCGATCCGGAGACGTTGTCGCCTTGTACA 400

410 420 430 440

TGGAGAACTCGGTCGAGTTTGTGGCCGCGTGGATGGGACT 440
CGCAAAAATCGGAGTTGTAACGGCTTGGATCAACTCGAAT 480
TTGAAAAGAGAGCAACTTGTTCATTGTATCACTGCGAGCA 520
AGACAAAGGCGATTATCACAAGTGTAACACTTCAGAATAT 560
TATGCTTGATGCTATCGATCAGAAGCTGTTTGATGTTGAG 600

610 620 630 640

GGAATTGAGGTTTACTCTGTCGGAGAGCCCAAGAAGAATT 640
CTGGATTCAAGAATCTCAAGAAGAAGTTGGATGCTCAAAT 680
TACTACGGAACCAAAGACCCTTGACATAGTAGATTTTAAA 720
AGTATTCTTTGCTTCATCTATACAAGTGGTACTACTGGAA 760
TGCCAAAAGCCGCTGTCATGAAGCACTTCAGATATTACTC 800

810 820 830 840

GATTGCCGTTGGAGCCGCAAAATCATTCGGAATCCGCCCT 840
TCTGATCGTATGTACGTCTCGATGCCAATTTATCACACTG 880
CAGCTGGAATTCTTGGAGTTGGGCAAGCTCTGTTGGGTGG 920
ATCATCGTGTGTCATTAGAAAAAAATTCTCGGCTAGCAAC 960
TTTTGGAGGGATTGTGTAAAGTATGATTGTACAGTTTCAC 1000

1010 1020 1030 1040

AATACATTGGAGAGATTTGTCGGTACTTGTTGGCTCAGCC 1040
AGTTGTGGAAGAGGAATCCAGGCATAGAATGAGATTGTTG 1080
GTTGGAAACGGACTCCGTGCTGAAATCTGGCAACCATTTG 1120
TAGATCGATTCCGTGTCAGAATTGGAGAACTTTATGGTTC 1160
AACTGAAGGAACTTCATCTCTCGTGAACATTGACGGACAT 1200

Fig. 76A

ceFATPa coding only.DNA

```
        1210        1220        1230        1240
GTCGGAGCTTGCGGATTCTTGCCAATATCCCCATTAACAA 1240
AGAAAATGCATCCGGTTCGATTAATTAAGGTTGATGATGT 1280
CACTGGAGAAGCAATCCGAACTTCCGATGGACTTTGCATT 1320
GCATGTAATCCAGGAGAGTCTGGAGCAATGGTGTCGACGA 1360
TCAGAAAAAATAATCCATTATTGCAATTCGAGGGATATCT 1400
        1410        1420        1430        1440
GAATAAGAAGGAAACGAATAAAAAAGATTATCAGAGATGTC 1440
TTCGCAAAGGGAGATAGTTGCTTTTTTGACTGGAGATCTTC 1480
TTCATTGGGATCGTCTTGGTTATGTATATTTCAAGGATCG 1520
TACTGGAGATACTTTCCGTTGGAAGGGAGAGAATGTGTCG 1560
ACTACTGAAGTCGAGGCAATTCTTCATCCAATTACTGGAT 1600
        1610        1620        1630        1640
TGTCTGATGCAACTGTTTATGGTGTAGAGGTTCCTCAAAG 1640
AGAGGGAAGAGTTGGAATGGCGTCAGTTGTTCGAGTTGTA 1680
TCGCATGAGGAAGATGAAACTCAATTTGTTCATAGAGTTG 1720
GAGCAAGACTTGCCTCTTCGCTTACCAGCTACGCGATTCC 1760
TCAGTTTATGCGAATTTGTCAGGATGTTGAGAAAACAGGT 1800
        1810        1820        1830        1840
ACATTCAAACTTGTGAAGACGAATCTACAACGATTAGGTA 1840
TCATGGATGCTCCTTCAGATTCAATTTACATCTACAATTC 1880
TGAAAATCGCAATTTTGTGCCGTTCGACAATGATTTGAGG 1920
TGCAAGGTCTCACTGGGAAGTTATCCATTTTAA 1953
```

Fig. 76B

ceFATPa coding only.protein

```
          10        20        30        40
          |         |         |         |
MKLEELVTVMLLTVAVIAQNLPIGVILAGVLILYITVVHG 40
DFIYRSYLTLNRDLTGLALIIEVKIDLWWRLHQNKGIHEL 80
FLDIVKKNPNKPAMIDIETNTTETYAEFNAHCNRYANYFQ 120
GLGYRSGDVVALYMENSVEFVAAWMGLAKIGVVTAWINSN 160
LKREQLVHCITASKTKAIITSVTLQNIMLDAIDQKLFDVE 200
          210       220       230       240
          |         |         |         |
GIEVYSVGEPKKNSGFKNLKKKLDAQITTEPKTLDIVDFK 240
SILCFIYTSGTTGMPKAAVMKHFRYYSIAVGAAKSFGIRP 280
SDRMYVSMPIYHTAAGILGVGQALLGGSSCVIRKKFSASN 320
FWRDCVKYDCTVSQYIGEICRYLLAQPVVEEESRHRMRLL 360
VGNGLRAEIWQPFVDRFRVRIGELYGSTEGTSSLVNIDGH 400
          410       420       430       440
          |         |         |         |
VGACGFLPISPLTKKMHPVRLIKVDDVTGEAIRTSDGLCI 440
ACNPGESGAMVSTIRKNNPLLQFEGYLNKKETNKKIIRDV 480
FAKGDSCFLTGDLLHWDRLGYVYFKDRTGDTFRWKGENVS 520
TTEVEAILHPITGLSDATVYGVEVPQREGRVGMASVVRVV 560
SHEEDETQFVHRVGARLASSLTSYAIPQFMRICQDVEKTG 600
          610       620       630       640
          |         |         |         |
TFKLVKTNLQRLGIMDAPSDSIYIYNSENRNFVPFDNDLR 640
CKVSLGSYPF. 651
```

Fig. 77

```
         10         20         30         40
         |          |          |          |
ATGAGGGAAATGCCGGACAGTCCCAAGTTTGCGTTAGTCA 40
CGTTTGTTGTGTATGCAGTGGTTTTGTACAATGTCAACAG 80
CGTTTTCTGGAAATTTGTATTCATCGGATATGTTGTATTT 120
AGGCTGCTTCGCACTGATTTTGGAAGAAGAGCACTTGCCA 160
CGTTACCTAGAGATTTTGCGGGACTGAAGCTCTTAATATC 200
        210        220        230        240
        |          |          |          |
GGTTAAGTCGACAATTCGTGGCTTGTTCAAGAAAGATCGC 240
CCAATTCATGAAATCTTTTTGAATCAGGTGAAACAGCATC 280
CAAACAAAGTGGCGATTATTGAAATTGAAAGTGGTAGGCA 320
GTTGACGTATCAAGAATTGAATGCGTTAGCTAATCAGTAT 360
GCTAACCTTTACGTGAGTGAAGGTTACAAAATGGGCGACG 400
        410        420        430        440
        |          |          |          |
TTGTCGCTTTGTTTATGGAAAATAGCATCGACTTCTTTGC 440
AATTTGGCTGGGACTTTCCAAGATTGGAGTCGTGTCGGCG 480
TTCATCAACTCAAACTTGAAGTTGGAGCCATTGGCACATT 520
CGATTAATGTTTCGAAGTGCAAATCATGCATTACCAATAT 560
CAATCTGTTGCCGATGTTCAAAGCCGCTCGTGAAAAGAAT 600
        610        620        630        640
        |          |          |          |
CTGATCAGTGACGAGATCCACGTGTTTCTGGCTGGAACTC 640
AGGTTGATGGACGTCATAGAAGTCTTCAGCAAGATCTCCA 680
TCTTTTCTCTGAGGATGAACCTCCAGTTATAGACGGACTC 720
AATTTTAGAAGCGTTCTGTGTTATATTTACACTTCCGGTA 760
CTACCGGAAATCCAAAGCCAGCCGTCATTAAACACTTCCG 800
        810        820        830        840
        |          |          |          |
TTACTTCTGGATTGCGATGGGAGCAGGAAAAGCATTTGGA 840
ATTAATAAGTCAGACGTTGTGTACATTACGATGCCAATGT 880
ATCACTCTGCCGCCGGTATCATGGGTATTGGATCATTAAT 920
TGCATTCGGGTCGACCGCTGTTATTAGGAAAAAGTTTTCG 960
GCAAGCAACTTCTGGAAAGATTGCGTCAAGTACAACGTCA 1000
        1010       1020       1030       1040
        |          |          |          |
CAGCGACACAGTACATTGGAGAAATCTGCAGGTATCTTCT 1040
GGCAGCGAATCCATGTCCTGAAGAGAAACAACACAACGTG 1080
CGATTGATGTGGGGAAATGGTTTGAGAGGACAAATTTGGA 1120
AAGAGTTTGTAGGAAGATTTGGAATTAAGAAAATTGGAGA 1160
GTTGTACGGCTCAACAGAAGGAAACTCCAATATTGTTAAC 1200
```

Fig. 78A

```
        1210        1220        1230        1240
GTGGATAACCATGTTGGAGCTTGTGGATTCATGCCAATTT 1240
ATCCCCATATTGGATCCCTCTACCCAGTTCGACTTATTAA 1280
GGTTGATAGAGCCACTGGAGAGCTTGAACGTGATAAGAAC 1320
GGACTCTGTGTGCCGTGTGTGCCTGGTGAAACTGGGGAAA 1360
TGGTTGGCGTTATCAAGGAGAAAGATATTCTTCTAAAGTT 1400
        1410        1420        1430        1440
CGAAGGATATGTCAGCGAAGGGGATACTGCAAAGAAAATC 1440
TACAGAGATGTGTTCAAGCATGGAGATAAGGTGTTTGCAA 1480
GTGGAGATATTCTTCATTGGGATGATCTTGGATACTTGTA 1520
CTTTGTGGACCGTTGTGGAGACACTTTCCGTTGGAAAGGG 1560
GAGAACGTGTCAACTACTGAAGTTGAGGGAATTCTTCAGC 1600
        1610        1620        1630        1640
CTGTGATGGATGTGGAAGATGCAACTGTTTATGGAGTCAC 1640
TGTCGGTAAAATGGAGGGGCGTGCCGGAATGGCTGGTATT 1680
GTCGTCAAGGATGGAACGGATGTTGAGAAATTCATCGCCG 1720
ATATTACTTCTCGACTGACCGAAAATCTGGCGTCTTACGC 1760
AATCCCTGTTTTCATTCGGCTGTGCAAGGAAGTTGATCGA 1800
        1810        1820        1830        1840
ACCGGAACCTTCAAACTCAAGAAGACTGATCTTCAAAAAC 1840
AAGGTTACGACCTGGTTGCTTGTAAAGGAGACCCAATTTA 1880
CTACTGGTCAGCTGCAGAAAAATCCTACAAACCACTGACT 1920
GACAAAATGCAACAGGATATTGACACTGGTGTTTATGATC 1960
GCATTTAA 1968
```

Fig. 78B

ceFATPb coding only.protein

```
        10        20        30        40
MREMPDSPKFALVTFVVYAVVLYNVNSVFWKFVFIGYVVF 40
RLLRTDFGRRALATLPRDFAGLKLLISVKSTIRGLFKKDR 80
PIHEIFLNQVKQHPNKVAIIEIESGRQLTYQELNALANQY 120
ANLYVSEGYKMGDVVALFMENSIDFFAIWLGLSKIGVVSA 160
FINSNLKLEPLAHSINVSKCKSCITNINLLPMFKAAREKN 200
       210       220       230       240
LISDEIHVFLAGTQVDGRHRSLQQDLHLFSEDEPPVIDGL 240
NFRSVLCYIYTSGTTGNPKPAVIKHFRYFWIAMGAGKAFG 280
INKSDVVYITMPMYHSAAGIMGIGSLIAFGSTAVIRKKFS 320
ASNFWKDCVKYNVTATQYIGEICRYLLAANPCPEEKQHNV 360
RLMWGNGLRGQIWKEFVGRFGIKKIGELYGSTEGNSNIVN 400
       410       420       430       440
VDNHVGACGFMPIYPHIGSLYPVRLIKVDRATGELERDKN 440
GLCVPCVPGETGEMVGVIKEKDILLKFEGYVSEGDTAKKI 480
YRDVFKHGDKVFASGDILHWDDLGYLYFVDRCGDTFRWKG 520
ENVSTTEVEGILQPVMDVEDATVYGVTVGKMEGRAGMAGI 560
VVKDGTDVEKFIADITSRLTENLASYAIPVFIRLCKEVDR 600
       610       620       630       640
TGTFKLKKTDLQKQGYDLVACKGDPIYYWSAAEKSYKPLT 640
DKMQQDIDTGVYDRI. 656
```

Fig. 79

chFATP coding only.DNA

```
 /      10        20        30        40
ATGGCGTGTATGCATCAGGCTCAGCTATACAATGATCTAG 40
AGGAATTGCTAACTGGTCCATCAGTACCCATCGTTGCTGG 80
AGCTGCTGGAGCTGCAGCTCTCACTGCCTACATTAACGCC 120
AAATACCACATAGCCCATGATCTCAAGACCCTCGGTGGTG 160
GATTGACACAATCGTCCGAAGCGATTGATTTCATAAACCG 200
       210       220       230       240
CCGCGTCGCACAAAAGCGCGTCCTCACGCACCACATCTTC 240
CAGGAGCAGGTCCAAAAACAATCAAATCATCCCTTTCTTA 280
TCTTTGAGGGCAAGACATGGTCTTACAAGGAGTTCTCTGA 320
GGCATACACGAGGGTCGCGAACTGGCTGATTGATGAGCTG 360
GACGTACAAGTAGGGGAGATGGTCGCAATTGATGGCGGAA 400
       410       420       430       440
ATAGTGCAGAGCACCTGATGCTTTGGCTTGCACTTGATGC 440
AATCGGTGCGGCTACGAGTTTTTTGAACTGGAACCTGACA 480
GGGGCAGGGTTAATTCATTGCATAAAGCTATGCGAATGTC 520
GATTCGTTATCGCAGACATCGATATTAAAGCGAACATTGA 560
ACCGTGCCGTGGCGAACTGGAGGAGACGGGCATCAACATT 600
       610       620       630       640
CACTACTATGACCCATCCTTCATCTCATCGCTACCGAATA 640
ACACGCCAATTCCCGACAGCCGCACTGAGAACATTGAATT 680
AGATTCAGTACGAGGACTGATATACACATCTGGAACCACT 720
GGTCTACCTAAAGGCGTGTTTATAAGCACTGGCCGCGAGC 760
TTAGGACTGACTGGTCGATTTCAAAGTATCTAAATCTCAA 800
       810       820       830       840
GCCCACGGATCGAATGTATACATGTATGCCGCTCTACCAT 840
GCCGCTGCACACAGCCTCTGTACAGCATCAGTTATTCATG 880
GTGGAGGTACCGTGGTATTGAGCAGGAAATTCTCACACAA 920
GAAGTTCTGGCCTGAAGTTGTGGCTTCGGAAGCAAATATC 960
ATTCAGTACGTTGGTGAATTAGGTCGATATCTCCTGAATG 1000
      1010      1020      1030      1040
GTCCAAAGAGTCCTTACGACAGGGCCCATAAAGTCCAGAT 1040
GGCGTGGGGCAATGGCATGCGTCCAGACGTGTGGGAAGCG 1080
TTTCGTGAACGCTTCAACATACCAATTATTCATGAGCTCT 1120
ATGCCGCAACCGATGGGCTCGGGTCAATGACCAATCGTAA 1160
CGCGGGCCCTTTTACAGCAAACTGTATTGCGCTGCGAGGG 1200
```

Fig. 80A

```
        1210        1220        1230        1240
....|....|....|....|....|....|....|....|....|....|
CTGATCTGGCACTGGAAATTTCGAAATCAGGAAGTGCTGG 1240
TCAAGATGGATCTCGATACTGATGAGATCATGAGAGATCG 1280
CAATGGGTTTGCGATACGATGCGCTGTCAATGAACCTGGA 1320
CAGATGCTTTTTCGGCTGACACCCGAAACTCTGGCTGGTG 1360
CACCAAGCTACTACAACAACGAAACGGCCACACAGAGCAG 1400
        1410        1420        1430        1440
....|....|....|....|....|....|....|....|....|....|
GCGGATTACAGATGTGTTTCAAAAGGGTGACCTGTGGTTC 1440
AAGTCCGGTGACATGCTACGGCAAGACGCCGAAGGCCGCG 1480
TCTACTTTGTCGATCGACTAGGCGATACGTTCCGCTGGAA 1520
ATCCGAAAACGTTTCTACCAATGAAGTCGCGGACGTGATG 1560
GGCACATTTCCTCAGATTGCTGAAACGAATGTATACGGTG 1600
        1610        1620        1630        1640
....|....|....|....|....|....|....|....|....|....|
TCCTTGTGCCGGGTAACGATGGTCGAGTGCGCAGCCTCAA 1640
TTGTCATGGCAGACGGCGTGACAGAGTCGACATTCGCTTC 1680
GCTGCCCTTGCAAAGCACGCCCGAGATCGGTTACCGGGTT 1720
ATGCTGTACCACTGTTTCTGAGGGTAACTCCAGCACTTGA 1760
ATATACGGGCACATTAAAGATTCAGAAAGGACGCCTCAAG 1800
        1810        1820        1830        1840
....|....|....|....|....|....|....|....|....|....|
CAGGAAGGTATAGACCCAGATAAGATTTCCGGCGAAGATA 1840
AGTTATACTGGCTGCCGCCTGGTAGCGATATATATTTACC 1880
ATTTGGAAAGATGGAGTGGCAGGGAATTGTAGATAAGCGT 1920
ATACGGCTGTGA 1932
```

Fig. 80B

chFATP coding only.protein

```
         10        20        30        40
MACMHQAQLYNDLEELLTGPSVPIVAGAAGAAALTAYINA 40
KYHIAHDLKTLGGGLTQSSEAIDFINRRVAQKRVLTHHIF 80
QEQVQKQSNHPFLIFEGKTWSYKEFSEAYTRVANWLIDEL 120
DVQVGEMVAIDGGNSAEHLMLWLALDAIGAATSFLNWNLT 160
GAGLIHCIKLCECRFVIADIDIKANIEPCRGELEETGINI 200
        210       220       230       240
HYYDPSFISSLPNNTPIPDSRTENIELDSVRGLIYTSGTT 240
GLPKGVFISTGRELRTDWSISKYLNLKPTDRMYTCMPLYH 280
AAAHSLCTASVIHGGGTVVLSRKFSHKKFWPEVVASEANI 320
IQYVGELGRYLLNGPKSPYDRAHKVQMAWGNGMRPDVWEA 360
FRERFNIPIIHELYAATDGLGSMTNRNAGPFTANCIALRG 400
        410       420       430       440
LIWHWKFRNQEVLVKMDLDTDEIMRDRNGFAIRCAVNEPG 440
QMLFRLTPETLAGAPSYYNNETATQSRRITDVFQKGDLWF 480
KSGDMLRQDAEGRVYFVDRLGDTFRWKSENVSTNEVADVM 520
GTFPQIAETNVYGVLVPGNDGRVRSLNCHGRRRDRVDIRF 560
AALAKHARDRLPGYAVPLFLRVTPALEYTGTLKIQKGRLK 600
        610       620       630       640
QEGIDPDKISGEDKLYWLPPGSDIYLPFGKMEWQGIVDKR 640
IRL 643
```

Fig. 81

```
          10        20        30        40
CTTTACCATTCATCAGCTTCATTCTGCATTTTTAGCTTGA 40
CGGCAGCCGGGTCTACGCTGATCATCGGCCGCAAGTTCTC 80
CGCGAGAAACTTCATAAAGGAAGCGCGCGAGAACGACGCC 120
ACGGTCATCCAGTACGTGGGTGAGACCTTGCGATATCTGC 160
TCGCCACCCCCGGTGAAACCGATCCAGTTACTGGCGAAGA 200
         210       220       230       240
CCTGGACAAAAAGCACAATATTCGAGCAGTATACGGCAAC 240
GGGCTACGGCCGGATATCTGGAACCGCTTCAAGGAGCGCT 280
TCAACGTGCCGACGGTTGCCGAATTTTATGCTGCAACCGA 320
GAGCCCAGGCGGAACATGGAACTATTCAACAAATGACTTC 360
ACTGCCGGAGCCATTGGGCACACTGGCGTGCTTAGTGGAT 400
         410       420       430       440
GGCTTCTTGGACGCGGCCTTACTATTGTCGAGGTGGACCA 440
GGAATCACAGGAACCATGGCGCGATCCCCAAACCGGGTTC 480
TGCAAGCCGGTCCCGCGAGGCGAAGCAGGCGAGCTCCTGT 520
ATGCCATTGATCCGGCCGACCCGGGCGAGACCTTCCAGGG 560
CTACTACCGCAACTCCTTTAGAGCACACTGGCGGCCG 597
```

Fig. 82

aspergillus partial.protein

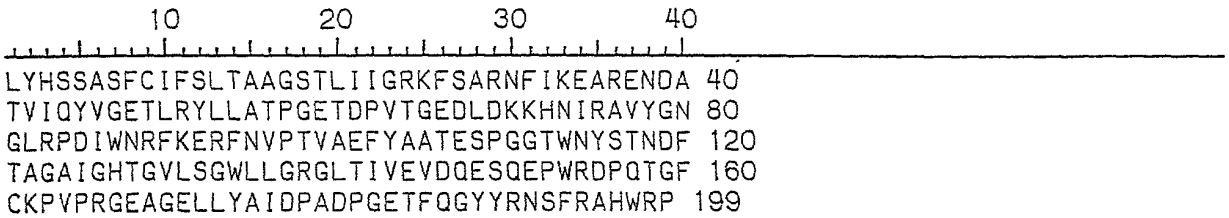


Fig. 83

```
        10         20         30         40
GCAAAGGCCGACGCGTGGCTGCGGACGGGTAACGTGATCA 40
GGGCGGACAACGAAGGGCGACTCTTCTTCCACGACCGGAT 80
CGGAGACACGTTCCGATGGAAGGGAGAGACNGTCAGCACA 120
CAAGAGGTCAGTTTGGTGCTCGGACGACACGACTCAATCA 160
AGGAGGCCAACGTGTACGGCGTGACGGTGCCGAACCACGA 200
        210        220        230        240
CGGGCGGGCCGGCTGCGCTGCGCTCACGCTATCAGACGCT 240
CTGGCGACTGAAAAGAAGCTGGGCGATGAGCTGCTAAAGG 280
GATTGGCTACTCACTCGTCGACTTCGCTTCCCAAGTTTGC 320
GGTGCCGCAGTTCCTACGGGTGGTGCGCGGCGAGATGCAG 360
TCAACGGGCACCAACAAGCAACAGAAGCACGACCTGAGGG 400
        410        420        430        440
TGCAGGGTGTAGAGCCGGGCAAGGTGGGCGTAGACGAGGT 440
GTACTGGTTGCGGGGAGGGACATATGTACCATTCGGAACA 480
GAGGATTGGGATGGGTTGAAGAAGGGTCTTGTGAAGTTGT 520
GA 522
```

Fig. 84

mgFATP partial.protein

```
         10        20        30        40
AKADAWLRTGNVIRADNEGRLFFHDRIGDTFRWKGETVST 40
QEVSLVLGRHDSIKEANVYGVTVPNHDGRAGCAALTLSDA 80
LATEKKLGDELLKGLATHSSTSLPKFAVPQFLRVVRGEMQ 120
STGTNKQQKHDLRVQGVEPGKVGVDEVYWLRGGTYVPFGT 160
EDWDGLKKGLVKL 173
```

Fig. 85

```
        10         20         30         40
....|....|....|....|....|....|....|....|....|
ATGTCTCCCATACAGGTTGTTGTCTTTGCCTTGTCAAGGA 40
TTTTCCTGCTATTATTCAGACTTATCAAGCTAATTATAAC 80
CCCTATCCAGAAATCACTGGGTTATCTATTTGGTAATTAT 120
TTTGATGAATTAGACCGTAAATATAGATACAAGGAGGATT 160
GGTATATTATTCCTTACTTTTTGAAAAGCGTGTTTTGTTA 200
        210        220        230        240
....|....|....|....|....|....|....|....|....|
TATCATTGATGTGAGAAGACATAGGTTTCAAAACTGGTAC 240
TTATTTATTAAACAGGTCCAACAAAATGGTGACCATTTAG 280
CGATTAGTTACACCCGTCCCATGGCCGAAAAGGGAGAATT 320
TCAACTCGAAACCTTTACGTATATTGAAACTTATAACATA 360
GTGTTGAGATTGTCTCATATTTTGCATTTTGATTATAACG 400
        410        420        430        440
....|....|....|....|....|....|....|....|....|
TTCAGGCCGGTGACTACGTGGCAATCGATTGTACTAATAA 440
ACCTCTTTTCGTATTTTTATGGCTTTCTTTGTGGAACATT 480
GGGGCTATTCCAGCTTTTTTAAACTATAATACTAAAGGCA 520
CTCCGCTGGTTCACTCCCTAAAGATTTCCAATATTACGCA 560
GGTATTTATTGACCCTGATGCCAGTAATCCGATCAGAGAA 600
        610        620        630        640
....|....|....|....|....|....|....|....|....|
TCGGAAGAAGAAATCAAAAACGCACTTCCTGATGTTAAAT 640
TAAACTATCTTGAAGAACAAGACTTAATGCATGAACTTTT 680
AAATTCGCAATCACCGGAATTCTTACAACAAGACAACGTT 720
AGGACACCACTAGGCTTGACCGATTTTAAACCCTCTATGT 760
TAATTTATACATCTGGAACCACTGGTTTGCCTAAATCCGC 800
        810        820        830        840
....|....|....|....|....|....|....|....|....|
TATTATGTCTTGGAGAAAATCCTCCGTAGGTTGTCAAGTT 840
TTTGGTCATGTTTTACATATGACTAATGAAAGCACTGTGT 880
TCACAGCCATGCCATTGTTCCATTCAACTGCTGCCTTATT 920
AGGTGCGTGCGCCATTCTATCTCACGGTGGTTGCCTTGCG 960
TTATCGCATAAATTTTCTGCCAGTACATTTTGGAAGCAAG 1000
        1010       1020       1030       1040
....|....|....|....|....|....|....|....|....|
TTTATTTAACAGGAGCCACGCACATCCAATATGTCGGAGA 1040
AGTCTGTAGATACCTGTTACATACGCCAATTTCTAAGTAT 1080
GAAAAGATGCATAAGGTGAAGGTTGCTTATGGTAACGGGC 1120
TGAGACCTGACATCTGGCAGGACTTCAGGAAGAGGTTCAA 1160
CATAGAAGTTATTGGTGAATTCTATGCCGCAACTGAAGCT 1200
```

Fig. 86A

scFATP coding only.DNA

1210 1220 1230 1240

CCTTTTGCTACAACTACCTTCCAGAAAGGTGACTTTGGAA 1240
TTGGCGCATGTAGGAACTATGGTACTATAATTCAATGGTT 1280
TTTGTCATTCCAACAAACATTGGTAAGGATGGACCCAAAT 1320
GACGATTCCGTTATATATAGAAATTCCAAGGGTTTCTGCG 1360
AAGTGGCCCCTGTTGGCGAACCAGGAGAAATGTTAATGAG 1400

1410 1420 1430 1440

AATCTTTTTCCCTAAAAAACCAGAAACATCTTTTCAAGGT 1440
TATCTTGGTAATGCCAAGGAAACAAAGTCCAAAGTTGTGA 1480
GGGATGTCTTCAGACGTGGCGATGCTTGGTATAGATGTGG 1520
AGATTTATTAAAAGCGGACGAATATGGATTATGGTATTTC 1560
CTTGATAGAATGGGTGATACTTTCAGATGGAAATCTGAAA 1600

1610 1620 1630 1640

ATGTTTCCACTACTGAAGTAGAAGATCAGTTGACGGCCAG 1640
TAACAAAGAACAATATGCACAAGTTCTAGTTGTTGGTATT 1680
AAAGTACCTAAATATGAAGGTAGAGCTGGTTTTGCAGTTA 1720
TTAAACTAACTGACAACTCTCTTGACATCACTGCAAAGAC 1760
CAAATTATTAAATGATTCCTTGAGCCGGTTAAATCTACCG 1800

1810 1820 1830 1840

TCTTATGCTATGCCCCTATTTGTTAAATTTGTTGATGAAA 1840
TTAAAATGACAGATAACCTCATAAAATTTTGA 1872

Fig. 86B

scFATP coding only.protein

```
          10        20        30        40
MSPIQVVVFALSRIFLLLFRLIKLIITPIQKSLGYLFGNY 40
FDELDRKYRYKEDWYIIPYFLKSVFCYIIDVRRHRFQNWY 80
LFIKQVQQNGDHLAISYTRPMAEKGEFQLETFTYIETYNI 120
VLRLSHILHFDYNVQAGDYVAIDCTNKPLFVFLWLSLWNI 160
GAIPAFLNYNTKGTPLVHSLKISNITQVFIDPDASNPIRE 200
         210       220       230       240
SEEEIKNALPDVKLNYLEEQDLMHELLNSQSPEFLQQDNV 240
RTPLGLTDFKPSMLIYTSGTTGLPKSAIMSWRKSSVGCQV 280
FGHVLHMTNESTVFTAMPLFHSTAALLGACAILSHGGCLA 320
LSHKFSASTFWKQVYLTGATHIQYVGEVCRYLLHTPISKY 360
EKMHKVKVAYGNGLRPDIWQDFRKRFNIEVIGEFYAATEA 400
         410       420       430       440
PFATTTFQKGDFGIGACRNYGTIIQWFLSFQQTLVRMDPN 440
DDSVIYRNSKGFCEVAPVGEPGEMLMRIFFPKKPETSFQG 480
YLGNAKETKSKVVRDVFRRGDAWYRCGDLLKADEYGLWYF 520
LDRMGDTFRWKSENVSTTEVEDQLTASNKEQYAQVLVVGI 560
KVPKYEGRAGFAVIKLTDNSLDITAKTKLLNDSLSRLNLP 600
         610       620       630       640
SYAMPLFVKFVDEIKMTDNLIKF. 624
```

Fig. 87

```
      10         20         30         40
GTGTCCGATTACTACGGCGGCGCACACACAACGGTCAGGC 40
TGATCGACCTGGCAACTCGGATGCCGCGAGTGTTGGCGGA 80
CACGCCGGTGATTGTGCGTGGGGCAATGACCGGGCTGCTG 120
GCCCGGCCGAATTCCAAGGCGTCGATCGGCACGGTGTTCC 160
AGGACCGGGCCGCTCGCTACGGTGACCGAGTCTTCCTGAA 200
     210        220        230        240
ATTCGGCGATCAGCAGCTGACCTACCGCGACGCTAACGCC 240
ACCGCCAACCGGTACGCCGCGGTGTTGGCCGCCCGCGGCG 280
TCGGCCCCGGCGACGTCGTTGGCATCATGTTGCGTAACTC 320
ACCCAGCACAGTCTTGGCGATGCTGGCCACGGTCAAGTGC 360
GGCGCTATCGCCGGCATGCTCAACTACCACCAGCGCGGCG 400
     410        420        430        440
AGGTGTTGGCGCACAGCCTGGGTCTGCTGGACGCGAAGGT 440
ACTGATCGCAGAGTCCGACTTGGTCAGCGCCGTCGCCGAA 480
TGCGGCGCCTCGCGCGGCCGGGTAGCGGGCGACGTGCTGA 520
CCGTCGAGGACGTGGAGCGATTCGCCACAACGGCGCCCGC 560
CACCAACCCGGCGTCGGCGTCGGCGGTGCAAGCCAAAGAC 600
     610        620        630        640
ACCGCGTTCTACATCTTCACCTCGGGCACCACCGGATTTC 640
CCAAGGCCAGTGTCATGACGCATCATCGGTGGCTGCGGGC 680
GCTGGCCGTCTTCGGAGGGATGGGGCTGCGGCTGAAGGGT 720
TCCGACACGCTCTACAGCTGCCTGCCGCTGTACCACAACA 760
ACGCGTTAACGGTCGCGGTGTCGTCGGTGATCAATTCTGG 800
     810        820        830        840
GGCGACCCTGGCGCTGGGTAAGTCGTTTTCGGCGTCGCGG 840
TTCTGGGATGAGGTGATTGCCAACCGGGCGACGGCGTTCG 880
TCTACATCGGCGAAATCTGCCGTTATCTGCTCAACCAGCC 920
GGCCAAGCCGACCGACCGTGCCCACCAGGTGCGGGTGATC 960
TGCGGTAACGGGCTGCGGCCGGAGATCTGGGATGAGTTCA 1000
     1010       1020       1030       1040
CCACCCGCTTCGGGGTCGCGCGGGTGTGCGAGTTCTACGC 1040
CGCCAGCGAAGGCAACTCGGCCTTTATCAACATCTTCAAC 1080
GTGCCCAGGACCGCCGGGGTATCGCCGATGCCGCTTGCCT 1120
TTGTGGAATACGACCTGGACACCGGCGATCCGCTGCGGGA 1160
TGCGAGCGGGCGAGTGCGTCGGGTACCCGACGGTGAACCC 1200
```

Fig. 88A

```
        1210        1220        1230        1240
....|....|....|....|....|....|....|....|....|....|
GGCCTGTTGCTTAGCCGGGTCAACCGGCTGCAGCCGTTCG 1240
ACGGCTACACCGACCCGGTTGCCAGCGAAAAGAAGTTGGT 1280
GCGCAACGCTTTTCGAGATGGCGACTGTTGGTTCAACACC 1320
GGTGACGTGATGAGCCCGCAGGGCATGGGCCATGCCGCCT 1360
TCGTCGATCGGCTGGGCGACACCTTCCGCTGGAAGGGCGA 1400
        1410        1420        1430        1440
....|....|....|....|....|....|....|....|....|....|
GAATGTCGCCACCACTCAGGTCGAAGCGGCACTGGCCTCC 1440
GACCAGACCGTCGAGGAGTGCACGGTCTACGGCGTCCAGA 1480
TTCCGCGCACCGGCGGGCGCGCCGGAATGGCCGCGATCAC 1520
ACTGCGCGCTGGCGCCGAATTCGACGGCCAGGCGCTGGCC 1560
CGAACGGTTTACGGTCACTTGCCCGGCTATGCACTTCCGC 1600
        1610        1620        1630        1640
....|....|....|....|....|....|....|....|....|....|
TCTTTGTTCGGGTAGTGGGGTCGCTGGCGCACACCACGAC 1640
GTTCAAGAGTCGCAAGGTGGAGTTGCGCAACCAGGCCTAT 1680
GGCGCCGACATCGAGGATCCGCTGTACGTACTGGCCGGCC 1720
CGGACGAAGGATATGTGCCGTACTACGCCGAATACCCTGA 1760
GGAGGTTTCGCTCGGAAGGCGACCGCAGGGCTAG 1794
```

Fig. 88B

mtFATP coding only.protein

```
         10         20         30         40
         |          |          |          |
MSDYYGGAHTTVRLIDLATRMPRVLADTPVIVRGAMTGLL 40
ARPNSKASIGTVFQDRAARYGDRVFLKFGDQQLTYRDANA 80
TANRYAAVLAARGVGPGDVVGIMLRNSPSTVLAMLATVKC 120
GAIAGMLNYHQRGEVLAHSLGLLDAKVLIAESDLVSAVAE 160
CGASRGRVAGDVLTVEDVERFATTAPATNPASASAVQAKD 200
        210        220        230        240
        |          |          |          |
TAFYIFTSGTTGFPKASVMTHHRWLRALAVFGGMGLRLKG 240
SDTLYSCLPLYHNNALTVAVSSVINSGATLALGKSFSASR 280
FWDEVIANRATAFVYIGEICRYLLNQPAKPTDRAHQVRVI 320
CGNGLRPEIWDEFTTRFGVARVCEFYAASEGNSAFINIFN 360
VPRTAGVSPMPLAFVEYDLDTGDPLRDASGRVRRVPDGEP 400
        410        420        430        440
        |          |          |          |
GLLLSRVNRLQPFDGYTDPVASEKKLVRNAFRDGDCWFNT 440
GDVMSPQGMGHAAFVDRLGDTFRWKGENVATTQVEAALAS 480
DQTVEECTVYGVQIPRTGGRAGMAAITLRAGAEFDGQALA 520
RTVYGHLPGYALPLFVRVVGSLAHTTTFKSRKVELRNQAY 560
GADIEDPLYVLAGPDEGYVPYYAEYPEEVSLGRRPQG. 598
```

Fig. 89

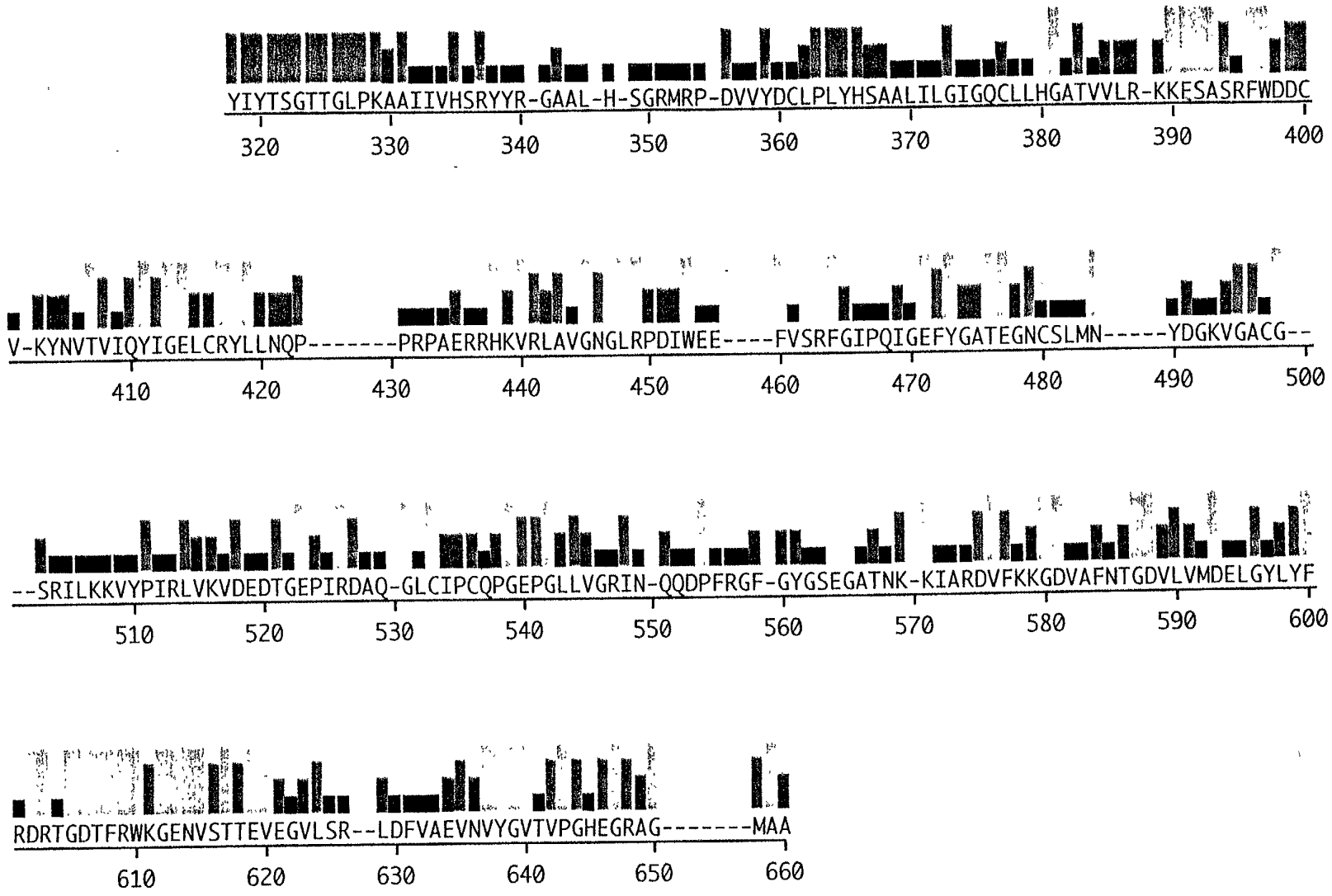

YIYTSGTTGLPKAAIIVHSRYYR-GAAL-H-SGRMRP-DVVYDCLPLYHSAALILGIGQCLLHGATVVLR-KKFSASRFWDDC
320 330 340 350 360 370 380 390 400
V-KYNVTVIQYIGELCRYLLNQP-------PRPAERRHKVRLAVGNGLRPDIWEE----FVSRFGIPQIGEFYGATEGNCSLMN-----YDGKVGACG--
410 420 430 440 450 460 470 480 490 500
--SRILKKVYPIRLVKVDEDTGEPIRDAQ-GLCIPCQPGEPGLLVGRIN-QQDPFRGF-GYGSEGATNK-KIARDVFKKGDVAFNTGDVLVMDELGYLYF
510 520 530 540 550 560 570 580 590 600
RDRTGDTFRWKGENVSTTEVEGVLSR--LDFVAEVNVYGVTVPGHEGRAG-------MAA
610 620 630 640 650 660

Figure 90

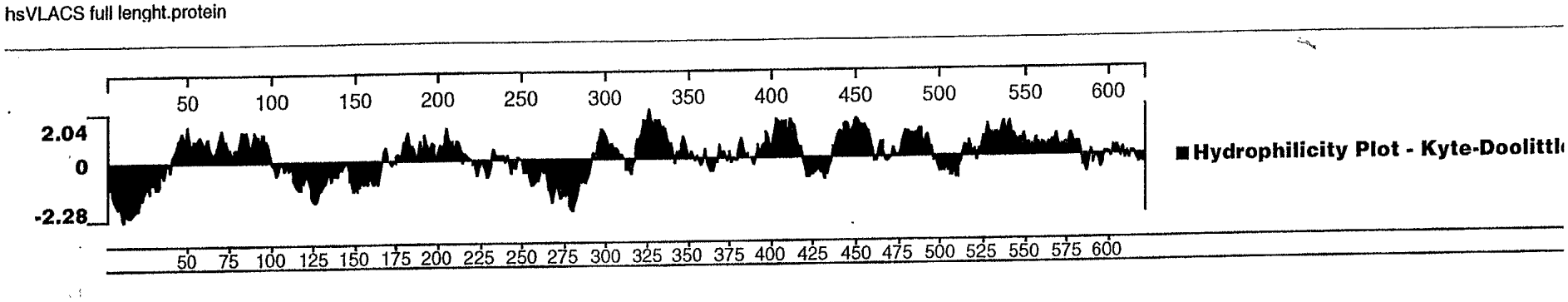


Figure 91

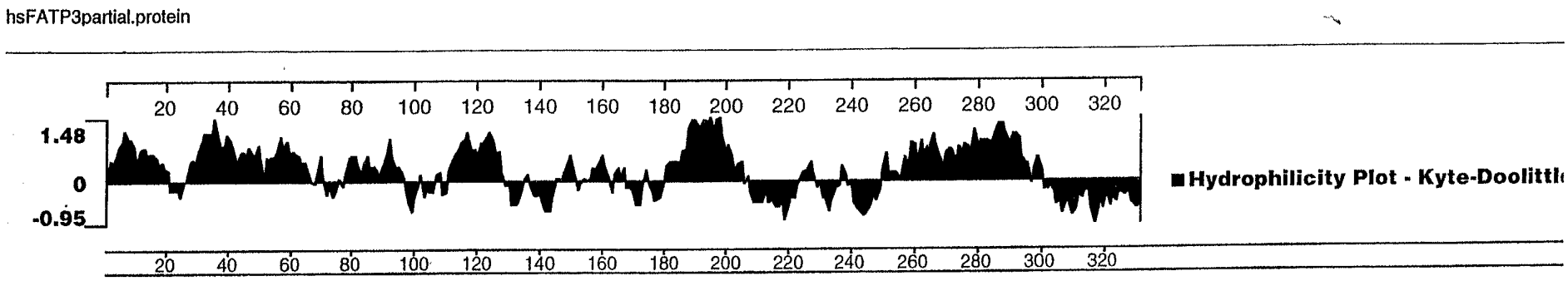


Figure 92

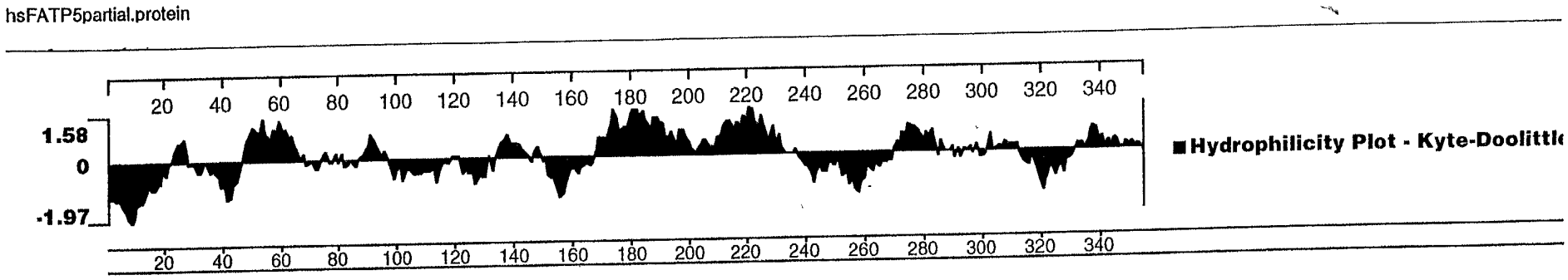

hsFATP5partial.protein
20 40 60 80 100 120 140 160 180 200 220 240 260 280 300 320 340
1.58
0
-1.97
20 40 60 80 100 120 140 160 180 200 220 240 260 280 300 320 340
Hydrophilicity Plot - Kyte-Doolittl

Figure 93

hsFATP3

```
1      cga ccc acg cgt ccg ggg atg ttt gcg agc ggc tgg aac cag acg gtg ccg ata gag gaa
1                          M   F   A   S   G   W   N   Q   T   V   P   I   E   E

61     gcg ggc tcc atg gct gcc ctc ctg ctg ctg ccc ctg ctg ctg ttg cta ccg ctg ctg ctg
15     A   G   S   M   A   A   L   L   L   L   P   L   L   L   L   L   P   L   L   L

121    ctg ctg aag cta cac ctc tgg ccg cag ttg cgc tgg ctt ccg gcg gac ttg gcc ttt gcg
35     L   L   K   L   H   L   W   P   Q   L   R   W   L   P   A   D   L   A   F   A

181    gtg cga gct ctg tgc tgc aaa agg gct ctt cga gct cgc gcc ctg gcc gcg gct gcc gcc
55     V   R   A   L   C   C   K   R   A   L   R   A   R   A   L   A   A   A   A   A

241    gac ccg gaa ggt ccc gag ggg ggc tgc agc ctg gcc tgg cgc ctc gcg gaa ctg gcc cag
75     D   P   E   G   P   E   G   G   C   S   L   A   W   R   L   A   E   L   A   Q

301    cag cgc gcc gcg cac acc ttt ctc att cac ggc tcg cgg cgc ttt agc tac tca gag gcg
95     Q   R   A   A   H   T   F   L   I   H   G   S   R   R   F   S   Y   S   E   A

361    gag cgc gag agt aac agg gct gca cgc gcc ttc cta cgt gcg cta ggc tgg gac tgg gga
115    E   R   E   S   N   R   A   A   R   A   F   L   R   A   L   G   W   D   W   G

421    ccc gac ggc ggc gac agc ggc gag ggg agc gct gga gaa ggc gag cgg gca gcg ccg gga
135    P   D   G   G   D   S   G   E   G   S   A   G   E   G   E   R   A   A   P   G

481    gcc gga gat gca gcg gcc gga agc ggc gcg gag ttt gcc gga ggg gac ggt gcc gcc aga
155    A   G   D   A   A   A   G   S   G   A   E   F   A   G   G   D   G   A   A   R

541    ggt gga gga gag ccc gcc gcc cct ctg tca cct gga gca act gtg gcg ctg ctc ctc ccc
175    G   G   G   E   P   A   A   P   L   S   P   G   A   T   V   A   L   L   L   P

601    gct ggc cca gag ttt ctg tgg ctc tgg ttc ggg ctg gcc aag gcc ggc ctg cgc act gcc
195    A   G   P   E   F   L   W   L   W   F   G   L   A   K   A   G   L   R   T   A

661    ttt gtg ccc acc gcc ctg cgc cgg ggc ccc ctg ctg cac tgc ctc cgc agc tgc ggc gcg
215    F   V   P   T   A   L   R   R   G   P   L   L   H   C   L   R   S   C   G   A

721    cgc gcg ctg gtg ctg gcg cca gag ttt ctg gag tcc ctg gag ccg gac ctg ccc gcc ctg
235    R   A   L   V   L   A   P   E   F   L   E   S   L   E   P   D   L   P   A   L

781    aga gcc atg ggg ctc cac ctg tgg gct gca ggc cca gga acc cac cct gct gga att agc
255    R   A   M   G   L   H   L   W   A   A   G   P   G   T   H   P   A   G   I   S

841    gat ttg ctg gct gaa gtg tcc gct gaa gtg gat ggg cca gtg cca gga tac ctc tct tcc
275    D   L   L   A   E   V   S   A   E   V   D   G   P   V   P   G   Y   L   S   S

901    ccc cag agc ata aca gac acg tgc ctg tac atc ttc acc tct ggc acc acg ggc ctc ccc
295    P   Q   S   I   T   D   T   C   L   Y   I   F   T   S   G   T   T   G   L   P

961    aag gct gct cgg atc agt cat ctg aag atc ctg caa tgc cag ggc ttc tat cag ctg tgt
315    K   A   A   R   I   S   H   L   K   I   L   Q   C   Q   G   F   Y   Q   L   C

1021   ggt gtc cac cag gaa gat gtg atc tac ctc gcc ctc cca ctc tac cac atg tcc ggt tcc
335    G   V   H   Q   E   D   V   I   Y   L   A   L   P   L   Y   H   M   S   G   S

1081   ctg ctg ggc atc gtg ggc tgc atg ggc att ggg gcc aca gtg gtg ctg aaa tcc aag ttc
355    L   L   G   I   V   G   C   M   G   I   G   A   T   V   V   L   K   S   K   F

1141   tcg gct ggt cag ttc tgg gaa gat tgc cag cag cac agg gtg acg gtg ttc cag tac att
375    S   A   G   Q   F   W   E   D   C   Q   Q   H   R   V   T   V   F   Q   Y   I

1201   ggg gag ctg tgc cga tac ctt gtc aac cag ccc ccg agc aag gca gaa cgt ggc cat aag
395    G   E   L   C   R   Y   L   V   N   Q   P   P   S   K   A   E   R   G   H   K
```

Figure 94A

```
1261  gtc cgg ctg gca gtg ggc agc ggg ctg cgc cca gat acc tgg gag cgt ttt gtg cgg cgc
415   V   R   L   A   V   G   S   G   L   R   P   D   T   W   E   R   F   V   R   R

1321  ttc ggg ccc ctg cag gtg ctg gag aca tat gga ctg aca gag ggc aac gtg gcc acc atc
435   F   G   P   L   Q   V   L   E   T   Y   G   L   T   E   G   N   V   A   T   I

1381  aac tac aca gga cag cgg ggc gct gtg ggg cgt gct tcc tgg ctt tac aag cat atc ttc
455   N   Y   T   G   Q   R   G   A   V   G   R   A   S   W   L   Y   K   H   I   F

1441  ccc ttc tcc ttg att cgc tat gat gtc acc aca gga gag cca att cgg gac ccc cag ggg
475   P   F   S   L   I   R   Y   D   V   T   T   G   E   P   I   R   D   P   Q   G

1501  cac tgt atg gcc aca tct cca ggt gag cca ggg ctg ctg gtg gcc ccg gta agc cag cag
495   H   C   M   A   T   S   P   G   E   P   G   L   L   V   A   P   V   S   Q   Q

1561  tcc cca ttc ctg ggc tat gct ggc ggg cca gag ctg gcc cag ggg aag ttg cta aag gat
515   S   P   F   L   G   Y   A   G   G   P   E   L   A   Q   G   K   L   L   K   D

1621  gtc ttc cgg cct ggg gat gtt ttc ttc aac act ggg gac ctg ctg gtc tgc gat gac caa
535   V   F   R   P   G   D   V   F   F   N   T   G   D   L   L   V   C   D   D   Q

1681  ggt ttt ctc cgc ttc cat gat cgt act gga gac acc ttc agg tgg aag ggg gag aat gtg
555   G   F   L   R   F   H   D   R   T   G   D   T   F   R   W   K   G   E   N   V

1741  gcc aca acc gag gtg gca gag gtc ttc gag gcc cta gat ttt ctt cag gag gtg aac gtc
575   A   T   T   E   V   A   E   V   F   E   A   L   D   F   L   Q   E   V   N   V

1801  tat gga gtc act gtg cca ggg cat gaa ggc agg gct gga atg gca gcc cta gtt ctg cgt
595   Y   G   V   T   V   P   G   H   E   G   R   A   G   M   A   A   L   V   L   R

1861  ccc ccc cac gct ttg gac ctt atg cag ctc tac acc cac gtg tct gag aac ttg cca cct
615   P   P   H   A   L   D   L   M   Q   L   Y   T   H   V   S   E   N   L   P   P

1921  tat gcc cgg ccc cga ttc ctc agg ctc cag gag tct ttg gcc acc aca gag acc ttc aaa
635   Y   A   R   P   R   F   L   R   L   Q   E   S   L   A   T   T   E   T   F   K

1981  cag cag aaa gtt cgg atg gca aat gag ggc ttc gac ccc agc acc ctg tct gac cca ctg
655   Q   Q   K   V   R   M   A   N   E   G   F   D   P   S   T   L   S   D   P   L

2041  tac gtt ctg gac cag gct gta ggt gcc tac ctg ccc ctc aca act gcc cgg tac agc gcc
675   Y   V   L   D   Q   A   V   G   A   Y   L   P   L   T   T   A   R   Y   S   A

2101  ctc ctg gca gga aac ctt cga atc tga gaa ctt cca cac ctg agg cac ctg aga gag gaa
695   L   L   A   G   N   L   R   I   *

2161  ctc tgt
```

Figure 94B

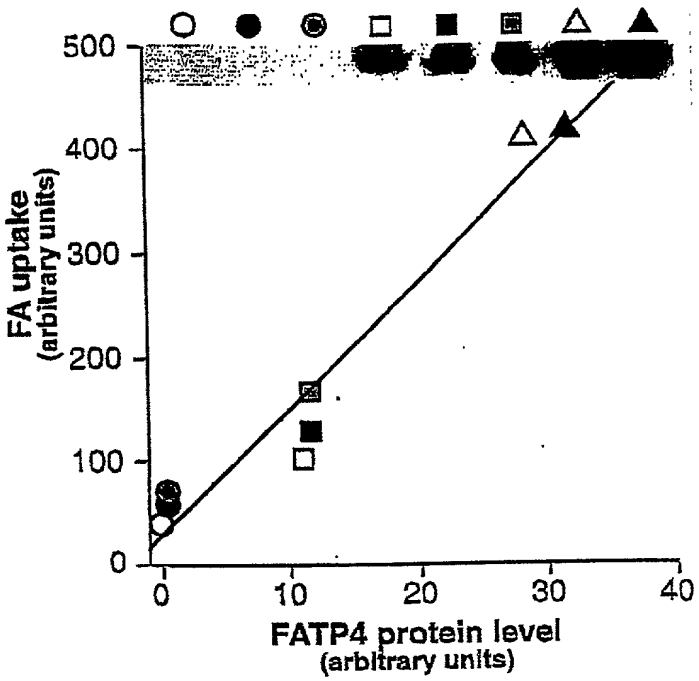

500
400
300
200
100
0
FA uptake
(arbitrary units)
0
10
20
30
40
FATP4 protein level
(arbitrary units)

Figure 95

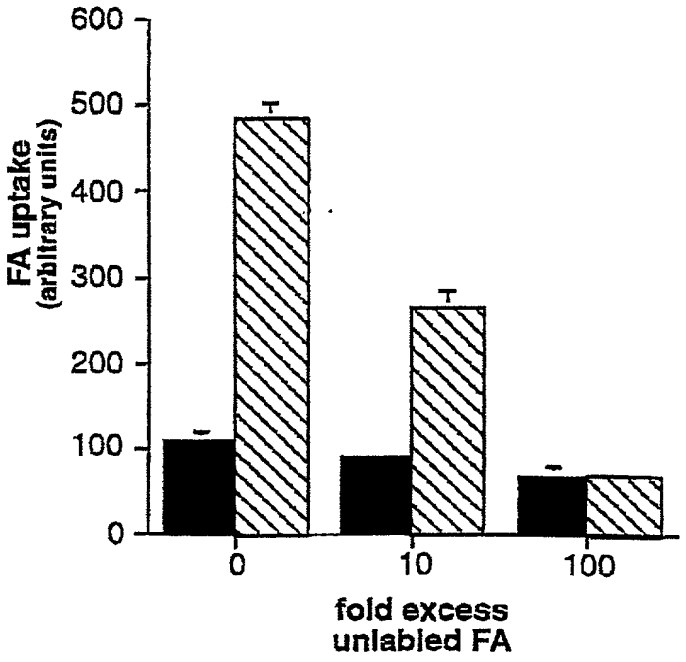

600
500
400
300
200
100
0
FA uptake
(arbitrary units)
0
10
100
fold excess
unlabled FA

Figure 96

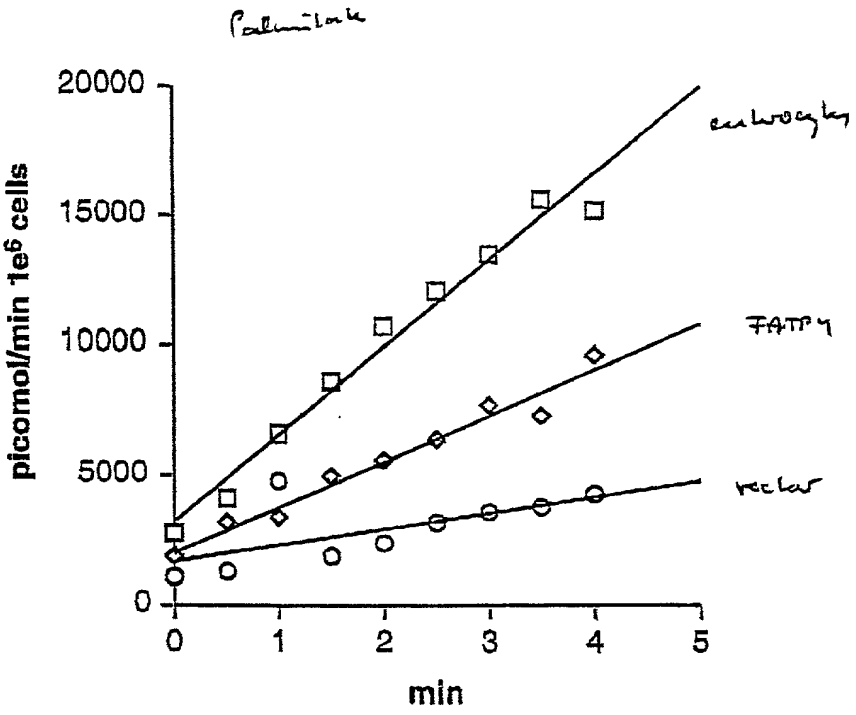

Palmitate
20000
15000
10000
5000
0
picomol/min 1e6 cells
0
1
2
3
4
5
min
FATP4

Figure 97

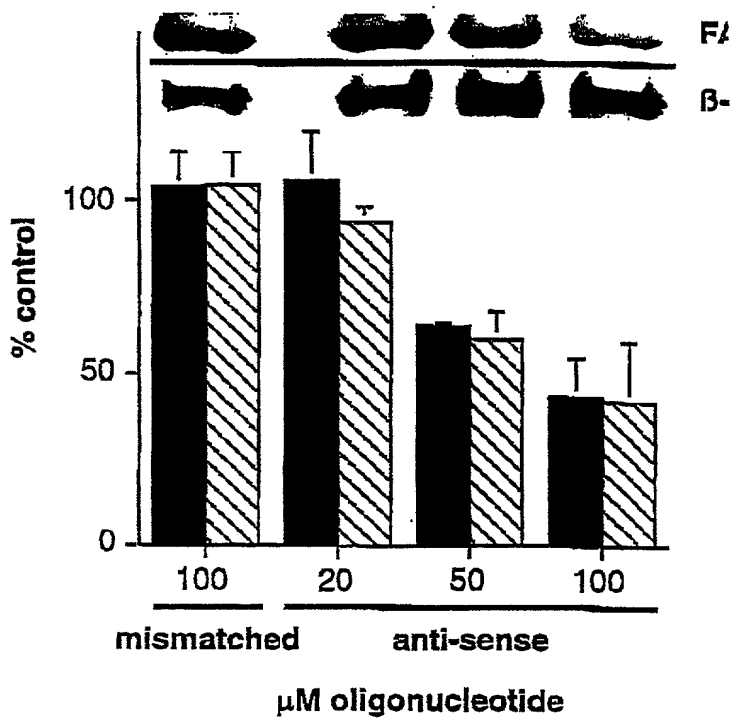

FA
β-
% control
100
50
0
100
20
50
100
mismatched
anti-sense
μM oligonucleotide

Figure 98

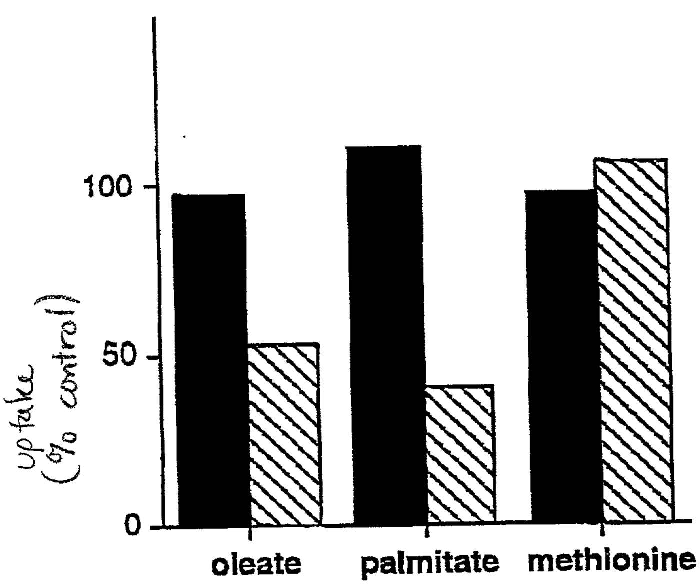


Figure 99